

5th ed.

MIM

Refining pre-polysomnography suspicion of Obstructive Sleep Apnea Syndrome: Logistic and Bayesian analysis of clinical factors

Liliana Patrícia Pinto Leite

MESTRADO EM
INFORMÁTICA MÉDICA
2º CICLO DE ESTUDOS

October | 2012



Refining pre-polysomnography suspicion of Obstructive Sleep Apnea Syndrome: Logistic and Bayesian analysis of clinical factors.

Liliana Patrícia Pinto Leite

MESTRADO EM
INFORMÁTICA MÉDICA
2º CICLO DE ESTUDOS

ORIENTADORES:

Pedro Pereira Rodrigues, Faculdade de Medicina da Universidade do Porto

Cristina Santos, Faculdade de Medicina da Universidade do Porto

Acknowledgements

This thesis is the culmination of one more academic step which means so much more than that. This Masters has given me the unique experience of meeting unbelievable people and acquiring knowledge that has allowed me to grow as a human being.

I want to thank all the people who were with me, in some way, during this process, either directly or indirectly.

To my parents, for all the lessons, love and sacrifices made throughout my life that really helped me to always overcome the difficulties.

To my mentors for instantly accepting this project with all the enthusiasm and dedication.

To Prof. Pedro Pereira Rodrigues for all the patience and teaching, for setting me objectives and challenges for which I thought I could not reach.

To Prof. Cristina Santos for the enthusiasm, interest and availability in the area.

To the Sleep Laboratory team of Vila Nova de Gaia Hospital Center for helping me on the work development, especially to my colleagues and friends of all hours, Dr. Gisela Fontes and Dr. Fernanda Azevedo for ensuring me the necessary schedule shifts that allowed me to attend the course and for helping me in the data collection.

To Pedro who encouraged me to attend the Masters, for all the love and patience and for making me always believe that it is worth fighting for and learning.

Resumo

Introdução: A Síndrome da Apneia Obstrutiva do Sono (SAOS) é uma doença que afecta 2-4% da população em todo o mundo. O método padrão para o diagnóstico de SAOS é a polissonografia (PSG), um exame caro, limitado às áreas urbanas e, consequentemente, com grandes listas de espera.

Objectivo: Definir um método auxiliar de diagnóstico que prioriza os pacientes, durante a consulta do sono, para a realização da PSG, de acordo com a probabilidade de diagnóstico de SAOS.

Métodos: Um estudo prospectivo foi realizado, incluindo pacientes adultos com suspeita de SAOS que realizaram PSG no laboratório do sono do Centro Hospitalar de Vila Nova de Gaia / Espinho. As variáveis estudadas foram definidas a partir de revisão de literatura e recolhidas durante a consulta. Foram colhidas duas amostras: uma coorte de treino para desenvolver os modelos e verificar a sua validade interna e externa através da validação cruzada (VC), e uma coorte de validação para verificar a aplicação dos modelos resultantes na prática clínica. Com as variáveis significativas obtidas na regressão logística (RL) univariada foram utilizadas duas técnicas diferentes para construir os modelos: RL múltipla e redes Bayesianas usando os classificadores Naïve Bayes (NB) e Tree Augmented Bayesian Network (TAN). A sensibilidade e especificidade foram analisadas para determinar a respectiva performance.

Resultados: Foram estudados 86 pacientes para construir os modelos, 52% dos quais com diagnóstico de SAOS. A RL univariada mostrou seis variáveis com influência significativa no resultado: sexo masculino (OR = 7,259, IC 95% = [1,096; 27,651]), índice de massa corporal (OR = 1,159, [1,030; 1,303]) circunferência do pescoço, (OR = 1,341, [1,159; 1,550]), circunferência abdominal (OR = 1,076, [1,025; 1,129]), apneias presenciadas (OR = 4,725, [1,772; 12,599]) e álcool antes de dormir (OR = 3,307, [1,350; 8,100]). Foram testados dois diferentes limites de sensibilidade. Com o intuito de obter 100% de sensibilidade foi utilizado um limite de 10% na RL obtido após a análise da curva ROC (AUC = 80% [70%, 89%]), 7% no NB e 2% no TAN, enquanto que se pretendermos 95% de sensibilidade, os limites obtidos foram de 25% na RL, 10% para o NB e 22% no TAN. A VC da RL estima que é robusto para ambos os limites: 98% para a sensibilidade e 11% para a especificidade e 89% -34%, respectivamente. No NB, usando 7% como limite, os resultados foram de 98% para a sensibilidade e de 18% para a especificidade, com o limite mais elevado (10%), os resultados foram de 93% de sensibilidade e de 30% de especificidade. No TAN, usando 2% como limite, os resultados foram de 88% de sensibilidade e 23% de

especificidade e com o limite de 22% foram de 84% sensibilidade e 25% de especificidade. Estes modelos foram testados numa segunda amostra comparável de 33 pacientes para avaliar o seu desempenho na prática clínica. Os resultados da RL apoiam as expectativas de ambos os limites (10% e 25%): 100% -0% e 88% -15%, respectivamente. Os limites de 7% e 2% usados para o NB e TAN respectivamente, obtiveram a mesma sensibilidade (94%), mas o TAN obteve melhor resultado relativamente à especificidade (7%) do que o NB (0%). Utilizando os limites mais elevados, 10% para o NB e de 22% para o TAN, os dois classificadores obtiveram mais uma vez a mesma sensibilidade (89%), mas o TAN revelou melhor resultado de especificidade (13%) do que o NB (7%).

Discussão: A circunferência do pescoço e apneias presenciadas fornecem informação suficiente para um modelo clínico com base nos resultados da RL. Se optarmos por redes Bayesianas devemos usar mais variáveis: sexo, índice de massa corporal, circunferência abdominal e álcool antes de dormir. Para ambos os modelos, utilizando os limites respectivos, podemos elaborar três níveis de prioridade dada a probabilidade de o paciente ter diagnóstico de SAOS: o grupo não prioritário, um nível intermédio e, por fim, um grupo de alta prioridade. Além destes resultados, o uso das redes revela duas principais vantagens que a RL tradicional não pode resolver. Primeiro, as redes Bayesianas podem lidar com informações em falta, e em segundo lugar, permitem uma representação gráfica que pode ser mais interessante para o médico. Consideramos que o uso destes modelos na consulta do sono pode ser uma ferramenta útil para a triagem de pacientes que realizem PSG e pode, eventualmente, ajudar a priorizar os doentes, permitindo talvez reduzir o número de PSG com resultado normal.

Palavras-chave: *factores de risco, Síndrome da Apneia Obstrutiva do Sono, diagnóstico, modelo clínico, redes Bayesianas, sensibilidade e especificidade.*

Abstract

Introduction: Obstructive Sleep Apnea (OSA) is a disease that affects 2-4% of the population around the world. The standard method for OSA diagnosis is polysomnography (PSG), an expensive exam, limited to urban areas and, consequently, with long waiting lists.

Aim: To define an auxiliary diagnostic method, that prioritizes patients during pre-polysomnography consultation, according to their probability of OSA diagnosis.

Methods: A prospective study was conducted, including adult patients with OSA suspicion that performed PSG at the Sleep Laboratory of Vila Nova de Gaia/Espinho Hospital Center. The studied variables were defined from literature review and collected during consultation. Two samples were collected: a training group to build the models and check internal and cross-validation (CV) and a validation group to check the resultant models in clinical practice. With the significant variables achieved with univariate logistic regression (LR) we used two different techniques, multiple LR and Bayesian networks classifiers- Naïve Bayes (NB) and Tree Augmented Bayesian network (TAN) - to build models that predicts OSA diagnosis. The sensitivity and specificity was analyzed to determine their performance.

Results: We studied 86 patients in order to build the models, 52% with OSA diagnosis. Univariate LR analysis showed six variables with significant influence on the outcome: male gender (OR=7.259, 95% CI=[1.096;27.651]), body mass index (OR=1.159, [1.030;1.303]), neck circumference (OR=1.341, [1.159;1.550]), abdominal circumference (OR=1.076, [1.025;1.129]), witnessed apneas (OR=4.725, [1.772;12.599]) and alcohol before sleep (OR=3.307, [1.350;8.100]). We tested two different cutoffs for sensitivity. Aiming 100% of sensitivity we used a 10% cutoff on LR achieved after a ROC curve analysis (AUC=80% [70%;89%]), 7% on NB and 2% on TAN while aiming 95% for sensitivity the cutoffs were 25% on LR, 10% for NB and 22% on TAN. The CV validation of LR model estimates that it was robust for both cutoffs (10% and 25%): 98%-11% and 89%-34%, respectively. On NB, using 7% as cutoff, the results were 98% for sensitivity and 18% for specificity and with the higher cutoff (10%) the results were 93% to sensitivity and 30% for specificity. On TAN, using 2% as cutoff, the results were 88% to sensitivity and 23% for specificity and to 22% were 84% to sensitivity and 25% for specificity. These models were tested on a separate comparable 33-patients cohort to analyze their performance on clinical practice. Results of LR supported the expectations for both thresholds: 100%-0% and 88%-15%,

respectively. The 7% (NB) and 2% (TAN) cutoffs obtained the same sensitivity (94%), but TAN achieved better results on specificity (7%) than NB (0%). Using the higher cutoffs of 10% on NB and 22% for TAN, the two classifiers obtained once again the same sensitivity (89%) but better results for specificity on TAN (13%) than in NB (7%).

Discussion: Neck circumference and witnessed apneas information suffices to a clinical model based on the LR results. If we use a BN we need more two variables: gender, body mass index, abdominal circumference and alcohol before sleep. For both models using the respective cutoffs we can provide three levels of priority given the probability of the patient having OSA diagnosis, non-priority group, an intermediate level and, finally, a priority group. Besides these results, the use of BN reveals two main advantages that traditional LR cannot solve. Firstly, BN can deal with missing information; second, the graphical representation can be more interesting to the physician. We consider that the use of these models on sleep consultation can be a helpful tool to monitor patients to perform PSG and eventually reduce the number of normal results PSG.

Key-words: *risk factors, obstructive sleep apnea, diagnosis, clinical model, Bayesian network, sensitivity and specificity.*

Preamble

The mysteries of sleep have always fascinated me which very naturally led me to a neurophysiology course. Luckily, my professional career began 8 years ago on the sleep laboratory of Vila Nova de Gaia/Espinho Hospital Center and the interest in the area has not decreased ever since.

The science of sleep is still very recent and we are still discovering new findings every year, new associations between sleep disorders and other diseases and, therefore, the increased interest in the area. Who has never had a sleepless night and all the unpleasant symptoms the next day? Sleep is a physiological need, so important as eating or drinking and we must give due attention in order to guarantee our well-being.

The evidence of the importance of sleep is reflected on long waiting list that we have in our laboratory, either for consultation or to perform polysomnography, which led me to wonder if we could find a way to manage all these cases, diagnosing and treating first the most severe cases, optimizing the available resources.

Index

Acknowledgements	iii
Resumo.....	iv
Abstract	vi
Preamble	viii
Index	ix
Abbreviations	x
Figure index	xi
Table index.....	xii
Thesis organization.....	xiii
Prior Dissemination	xiv
1. Introduction.....	1
2. Aim	5
3. Background	6
4. Material and Methods	28
5. Results	33
6. Discussion	43
7. Conclusions and recommendations	51
8. Future work	52
9. References	53
Attachments	59

Abbreviations

AASM: American Academy of Sleep Medicine

AC: Abdominal circumference

AHI: Apnea-hypopnea index

AUC: Area under the curve

BMI: Body mass index

BN: Bayesian networks

BQ: Berlin Questionnaire

CV: Cross-validation

DM: Data Mining

ECG: Electrocardiography

EEG: Electroencephalography

EMG: Electromyography

EOG: Electro-oculography

ESS: Epworth Sleepiness Scale

LR: Logistic regression

NB: Naïve Bayes

NC: Neck circumference

NREM: Non Rapid Eyes Movement

NPV: Negative predictive value

OSA: Obstructive Sleep Apnea

OR: Odds ratios

PM: Portable Monitors

PPV: Positive predictive value

PSG: Polysomnography

RDI: Respiratory disturbance index

REM: Rapid Eyes Movement

ROC: Receiver Operating Characteristic

TAN: Tree augmented Bayesian network

WA: Witnessed apneas

Figure index

Figure 1: Naive Bayes networks.....	25
Figure 2: Tree augmented Bayesian network.....	25
Figure 3: Flow diagram for inclusion of patients in the study	33
Figure 4: ROC curve of the model.....	36
Figure 5: Bayesian network and the conditional probabilities	37
Figure 6: Naive Bayes classifier.....	38
Figure 7: Marginal probabilities of Naïve Bayes classifier to the presence of OSA	39
Figure 8: Tree augmented Bayesian network.....	39
Figure 9: Marginal probabilities of TAN	40
Figure 10: Flow diagram for inclusion of patients in the second comparable study.....	41
Figure 11: Inference using NB with new evidence: female, without WA nor alcohol consumption before sleep with AC, NC and BMI normal.....	46
Figure 12: Inference using TAN with new evidence: with new evidence: female, without WA nor alcohol consumption before sleep with AC, NC and BMI normal.	46
Figure 13: Inference using NB with missing information of WA.....	47
Figure 14: Inference using TAN with missing information of WA.....	48
Figure 15: Inference using NB with missing information of WA, AC and NC.....	48
Figure 16: Inference using TAN with missing information of WA, AC and NC.....	49

Table index

Table 1: Criteria for score sleep stages.....	8
Table 2: Scoring of respiratory events.....	9
Table 3: Types of sleep studies used in sleep laboratories.....	11
Table 4: Variables collected in the studied patients	29
Table 5: Parameters used to recode continuous variables BMI, NC and AC into dichotomous.....	31
Table 6: Description and odds ratios for the 33 studied variables	34
Table 7: Characteristics of two samples.....	41
Table 8: Results of LR and BN on the train and test sample.....	42

Thesis organization

In chapter 1 (Introduction) we make an introduction to the problem that motivates this thesis and explains the solutions tested and the various methods developed to screen patients to perform polysomnography (PSG).

In Chapter 2 (Aim) we set the main objective which is to define an auxiliary diagnostic method that can support the decision to perform polysomnography, or prioritize the waiting list to polysomnography.

In Chapter 3 (Background) various definitions are quoted to help in the thesis perception. Concepts like sleep apnea, risk factors to OSA, logistic regression and Bayesian networks are explain.

In Chapter 4 (Methods) we present the methodology used to develop and evaluate the models, including variables collected, the selection criteria used to select the two samples, the parameters used to construct two models based in logistic regression and Bayesian networks and finally the models validation.

In Chapter 5 (Results) we present the results of the samples selection and the performance of the models in both samples.

In Chapter 6 (Discussion) the risk factors for obstructive sleep apnea (OSA) achieved on the univariate logistic regression are compared to the related on literature. The performance, advantages and disadvantages of the models are discussed.

In Chapter 7 (Conclusions and recommendations) we expose a brief conclusion of our results and their explanations, as well as the main recommendations to apply these models on clinical practice

In Chapter 8 (Future work) we explain the various possibilities to create a decision support system as in a sleep laboratory, based on ours models.

Prior Dissemination

The investigation protocol to develop this thesis with the provisional title "Data mining as an auxiliary diagnostic for the Syndrome of Obstructive Sleep Apnea: Is it possible to reduce the number of unnecessary polysomnographies?", was shown on 4th Symposium on Medical Informatics, October 2011, Porto, Portugal

The preliminary results was presented on the Intelligent Data Analysis meeting organized by the Health Information and Decision Sciences department, Faculty of Medicine, University of Porto, Portugal on 25th of January of 2012.

1. Introduction

The syndrome of Obstructive Sleep Apnea (OSA) is a disease that affects approximately 4% of men and 2% of women worldwide but is still underestimated and underdiagnosed (Al Lawati, Patel, & Ayas, 2009; Jennum & Riha, 2009; Madani & Madani, 2009; T. Young, Evans, Finn, & Palta, 1997). It is characterized by episodes of breathing cessation (apnea) or reduction in airflow (hypopnea) during sleep for at least 10 seconds as a result of a upper airway collapse (Al Lawati et al., 2009; Iber C, 2007; Rechtschaffen, 1968; Redline et al., 2007; Silber et al., 2007). The severity of OSA is associated with the apnea-hypopnea index (AHI), documented during sleep, which can be divided into mild ($5 \leq \text{AHI} < 15$), moderate ($15 \leq \text{AHI} < 30$) and severe ($\text{AHI} \geq 30$) ("Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force," 1999). The standard method for assessing this index, and therefore defining the OSA diagnosis, is polysomnography (PSG). However, it is time-consuming, expensive and relatively limited to urban areas which, consequently, originates high waiting lists (Sun, Chiu, Chuang, & Liu, 2010).

For a correct diagnosis, it is important to first determine the factors associated with the disease, and then use them to calculate the probability of the presence of OSA. According to the literature, the risk factors associated with OSA are age, gender and body mass index (BMI). However, there is no consensus on the weight of these factors in the prediction (Al Lawati et al., 2009; Davies, Ali, & Stradling, 1992; Doghramji, 2008; Hoffstein & Szalai, 1993; Kapur, 2010; Kohler, 2009; Manber & Armitage, 1999; T. Young et al., 1997; Terry Young et al., 2002; T. Young, Skatrud, & Peppard, 2004). Some authors referred other features such as neck circumference (NC), witnessed apneas or diurnal somnolence as important risk factors too (Davies et al., 1992; Pouliot, Peters, Neufeld, & Kryger, 1997). The path for a consensus is still undetermined.

In Portugal, patients are referred by the primary care physician to a sleep consult, and then the sleep expert physicians decide the need to perform polysomnography. Although patients are screened by the physicians, based on clinical factors, the specificity of the entire process is rather low (48% of PSG performed in 2010, in our sleep laboratory, resulted negative for OSA, from which 75% had a completely normal result for sleep disorders) which, together with the limited availability of the service, yields long waiting lists both for consultation and to perform PSG. This problem is also prevalent in other sleep laboratories and several studies have been conducted to define the most important factors to determine

the probability of having OSA, and thereby reduce and optimize the number of patients that realize PSG, assigning different priority to patients (Al Lawati et al., 2009; Davies et al., 1992; Dixon, Schachter, & O'Brien, 2003; Flemons, Whitelaw, Brant, & Remmers, 1994; Gurubhagavatula, Maislin, & Pack, 2001; Hoffstein & Szalai, 1993; Maislin et al., 1995; Pouliot et al., 1997; Rodsutti, Hensley, Thakkestian, D'Este, & Attia, 2004; Viner, Szalai, & Hoffstein, 1991; Terry Young et al., 2002).

To identify more quickly OSA patients and possibly reduce the number of PSGs, some procedures have been adopted like the use of Portable Monitors (PM) and prediction models, but these don't show capable to stop the tendency of increased waiting lists.

PM are a useful tool in cases of patients without comorbid conditions or medical disorders, with a higher probability of moderate or severe OSA. Otherwise, this method tends to underestimate severity of OSA, because don't allow determine sleep efficiency, and so, PSG have to be performed on the most cases (Collop et al., 2007).

Prediction models were built based on questionnaires and prediction methods to screen patients with a higher probability of OSA diagnosis (Flemons et al., 1994; Gurubhagavatula et al., 2001; Kaimakamis, Bratsas, Sichletidis, Karvounis, & Maglaveras, 2009; Kwiatkowska, Atkins, Ayas, & Ryan, 2007; Maislin et al., 1995; Pouliot et al., 1997; Sun et al., 2010; Viner et al., 1991). Traditionally, these models consisted in simple decision rules, the prognostic score and classification of patients into different risk categories. This score is often based on the combination of clinical variables and has been built for the general population, as well as for specific groups. These models can be an alternative method in a sleep consultation to help in the clinical decision to perform PSG. But, to construct clinical decision rules, we have to check some characteristics of the models to validate their use in clinical practice (Kononenko, 2001). The main limitation is sensitivity. These models need a high sensitivity, as false negatives should be avoided, to prevent excluding a patient with moderate or severe OSA from performing PSG. No study founded on literature in was fitted for 100% sensitivity.

In Portugal, we found one study that tried to implement a screening tool for OSA. Vaz et al. used the Berlin Questionnaire (BQ), one of the most recognized screening tool, to screen patients with OSA in a sleep breathing clinic (Vaz et al., 2011). It includes 10 items organized in 3 categories concerning snoring and witnessed apneas (5 items), daytime sleepiness (4 items) and high blood pressure /obesity (1 item). Patients are also asked to provide information on age, gender, weight, height, neck circumference and ethnicity. Predetermination of high or lower risk for OSA is based on responses to each category of items. The authors achieve a sensitivity of 65.2 % and specificity of 80%, what reveals a good discrimination but poor performance in OSA identification. These results are similar to other studies that have different performances and don't reveal BQ as an alternative to screen patients (Ahmadi, Chung, Gibbs, & Shapiro,

2008; Chung et al., 2008; Gami et al., 2004; Gus et al., 2008; Netzer, Stoohs, Netzer, Clark, & Strohl, 1999; Weinreich, Plein, Teschler, Resler, & Teschler, 2006).

Another frequent problem in the application of the models is the lack of internal and/or external validation of the results (Davies et al., 1992; Flemons et al., 1994; Hoffstein & Szalai, 1993; Viner et al., 1991; Terry Young et al., 2002) which compromises their application. Some studies don't show the regression parameters (Hoffstein & Szalai, 1993; Terry Young et al., 2002); other use factors that are subjective and may lead to lower measurement reliability (Hoffstein & Szalai, 1993). Another point that limits their application is the choice of variables used in the construction of the models. As they are based on clinical variables, the missing of an important one compromises their results and, consequently, their validation. Neck circumference is one of the most significant clinical variable to identify patients with OSA according to some studies (Davies et al., 1992) but other studies don't use this measure to construct their models (Pouliot et al., 1997; Rodsutti et al., 2004; Viner et al., 1991). Furthermore, the American Academy of Sleep Medicine (AASM) recommends the use of the 5 events per hour cutoff to distinguish patients with OSA from those where OSA is absent ("Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force," 1999). Some studies used different cutoffs to describe OSA (10, 15, 20 or 30) which is questionable (Flemons et al., 1994; Hoffstein & Szalai, 1993; Maislin et al., 1995; Pouliot et al., 1997; Viner et al., 1991) and can introduce a large number of false negatives and difficulties in their comparison.

Today, these models are generated by artificial intelligence, using decision trees, neural networks, support vector machines and Bayesian networks (BN) (Lee & Abbott, 2003; van Gerven, Taal, & Lucas, 2008). To be useful the models must have certain characteristics, such as good performance, good ability to handle data entry errors or omissions, transparency of diagnostic knowledge, ability to explain decisions, and the algorithm is able to reduce the number of tests needed for make a reliable diagnosis (Kononenko, 2001)

Data mining involves the extraction of information, whose goal is to discover facts and / or unknown or hidden patterns in a database and extensive inference rules to predict trends. This process is based on combinations of machine learning and statistical analysis (Lavrac, 2001; Lee & Abbott, 2003; P. Lucas, 2004; Mitchell, 1997; van Gerven et al., 2008). The data mining tools have been used in sleep medicine to create models alternative of those based in logistic regression like decision trees applied to PSG signals, genetic algorithm to screen patients with moderate or severe OSA.

Bayesian networks, in particular, have been used in medical domain in some areas with high performance like in diagnoses of pneumonia and breast cancer, classification of cytological findings, classification, prediction of patient compliance to medication, prediction of clinician compliance to medical practice guidelines, prognosis of head injuries, determination of the risk factors of obesity, and pattern recognition

in narrative clinical reports (Aronsky & Haug, 2000; E. Burnside, Rubin, & Shachter, 2000; E. S. Burnside, 2005; Hamilton et al., 1995; Lee & Abbott, 2003; Montironi, Bartels, Thompson, Scarpelli, & Hamilton, 1995; Sakellaropoulos & Nikiforidis, 1999; Taktak, Ajmi Nabli, Ben Othmen, Mtiraoui, & Ben Hadj Hamida, 2011). As BN are a power data mining tool, we choose this method to create a model to screen patients with OSA.

2. Aim

The main objective of this work is to define an auxiliary diagnostic method that can support the decision to perform polysomnography, or prioritize the waiting list to polysomnography, in patients recommended by a general practitioner, suspected of having OSA, specifically:

- Prioritize patients recommended for PSG;
- Reduce the number of “unnecessary polysomnographies” (increase specificity) or give a higher priority in the waiting list for polysomnography to patients with more chances of OSA diagnosis;
- Avoid the recommendation “unnecessary polysomnography” to OSA cases (avoid false negatives);
- Produce effective models for use in clinical practice;
- Expand the graphical interpretation of results.

3. Background

The upper airway includes the extrathoracic trachea, larynx, pharynx, nose and is separated into three regions: the nasopharynx, which is defined from the nasal turbinates to the hard palate; the oropharynx, subdivided into the retro palatal region ; and the hypopharynx (Kryger, 2005).

To study obstructive sleep apnea (OSA) we will focus on the pharyngeal airway, specifically the retropalatal retroglossal regions because is the site of upper airway closure or narrowing during sleep in the majority of patients with OSA. This part is a conduit for airflow connecting the nose with the larynx, pharyngeal patency is critical. With the exception of the two ends of the respiratory airway tract (the nares and the small intrapulmonary airways), the pharynx is the only collapsible segment of the respiratory tract (Kryger, 2005).

The sleep state is associated with a decrease in motor output to pharyngeal muscles. When this occurs against the background of upper airway anatomic abnormalities, severe narrowing or closure of pharyngeal airway can occur.

3.1. Pathogenesis of OSA

The pathogenesis of OSA may be explained by some factors: alterations of upper-airway dilator muscle activity during sleep and his anatomy, lung volume, ventilatory control stability, sleep state stability and rostral fluid shifts (Kapur, 2010; Yaggi & Strohl, 2010). The relationship of this factors influence breath and depends on a balance of forces: forces that promote airway collapse and opposing forces that maintain upper airway patency (Yaggi & Strohl, 2010).

The balance of forces promoting airway collapse, like negative pressure of ventilation and extraluminal positive pressure, and forces to oppose these collapsing, activity of the pharyngeal dilator muscles (eg, genioglossus) and tensor palatine dilator muscles are tonically active, are usually maintained during sleep (Yaggi & Strohl, 2010) but, in patients with OSA, some of this controls are lost. Dilator muscle activity is controlled by genioglossus, the muscle that forms the majority of the body of the tongue (Kapur, 2010; Yaggi & Strohl, 2010), and is responsible for stiffen and dilate various regions of the airway. Any

alterations in muscle activity or lower end-expiratory lung volume, increases the tendency of the upper airway collapse (Kapur, 2010).

3.1.1. Ventilatory control stability

OSA causes great instability of ventilatory control system and, consequently, a higher loop gain (measure of the stability of a negative-feedback control system) because autonomous system have to response to the inputs generated by the upper-airway muscles (Kapur, 2010).

3.1.2. Stability of sleep

The sleep stability is affected by the number increased of arousals that occurs as a biological response to the hypoxemia, causing an increase in respiratory effort and accentuate the changes in ventilation. Neural respiratory control centers response changing the level of PaO_2 and PaCO_2 augmenting and perpetuating respiratory cycling (Kapur, 2010; Yaggi & Strohl, 2010).

3.1.3. Rostral fluid shifts

Fluid displacement from the legs caused by lower body positive pressure, by inflation of antishock trousers increases neck circumference, narrows the pharynx, and increases collapsibility in awake healthy subjects has been shown to reduce upper-airway size and increase collapsity (Kapur, 2010; Yaggi & Strohl, 2010).

3.2. Sleep stages and event scoring

The American Academy of Sleep Medicine (AASM) recommend the criteria scoring of sleep stages and the use of the terminology division into wakefulness, Non Rapid Eyes Movement (NREM) with 3 stages N1, N2 and N3, and REM (Rapid Eyes Movement)(Iber C, 2007; Silber et al., 2007). Table 1 summarizes the main characteristics of the five stages.

Table 1: Criteria for score sleep stages

Stages	Rules
Wakefulness	<p>Eye blinks at a frequency of 0.52Hz</p> <p>Reading eye movements</p> <p>Irregular conjugate rapid eye movements associated with normal or high chin muscle tone</p> <p>Epochs without discernible alpha rhythm</p>
Stage N1	<p>A. In subjects who generate alpha rhythm, score stage N1 if alpha rhythm is attenuated and replaced by low amplitude, mixed frequency activity for more than 50% of the epoch.</p> <p>B. In subjects who do not generate alpha rhythm, score stage N1 commencing with the earliest of any of the following phenomena:</p> <ol style="list-style-type: none"> 1) Activity in range of 4-7Hz with slowing of background frequencies by ≥ 1 Hz from those of stage W. 2) Vertex sharp waves. 3) Slow eye movements.
Stage N2:	<p>A. One or more K complexes unassociated with arousals</p> <p>B. One or more trains of sleep spindles</p>
Stage N3	<p>20% or more of an epoch consists of waves of 0.5-2 Hz frequencies with peak-to-peak amplitude of $>75 \mu\text{V}$ in the frontal derivation.</p>
Stage R (REM sleep)	<p>Presence of eye movements are and stage 2 absent or low amplitude mixed frequency EEG and persistently low chin EMG tone</p>
REM: Rapid Eyes Movement.	

According to the AASM, respiratory events can be divided into apnea or hypopnea (table 2).

Table 2: Scoring of respiratory events

Event	Criteria
Apnea	Drop in the peak thermal sensor excursion of at least 90% of baseline During at least 10 seconds At least 90% of event's meets the amplitude reduction criteria for apnea Obstructive: Continued or increased inspiratory effort during the apnea Central: Absence of inspiratory effort during the apnea Mixed: Absence of inspiratory effort in the initial portion of the event followed by resumption of inspiratory effort in the second portion
Hypopnea (Recommended rules)	The nasal pressure excursions drop by $\geq 30\%$ of baseline The duration of this drop is at least 10 seconds There is a $\geq 4\%$ desaturation from pre-event baseline or the event is associated with arousal At least 90% of event's meets the amplitude reduction criteria for hypopnea
Hypopnea (Alternative rules)	The nasal pressure excursions drop by $\geq 50\%$ of baseline The duration of this drop is at least 10 seconds There is a $\geq 3\%$ desaturation from pre-event baseline or the event is associated with arousal At least 90% of event's meets the amplitude reduction criteria for hypopnea
RERA	Sequence of breaths during at least 10 seconds characterized by increasing respiratory effort or flattening of the nasal pressure waveform leading to an arousal from sleep when does not meet the criteria for apnea or hypopnea
RERA: Respiratory effort-related arousal	

3.3. OSA diagnosis

The diagnosis of OSA has to follow the criteria A or B plus C (*Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force*, 1999):

A. Excessive daytime sleepiness that is not better explained by other factors;

B. Two or more of the following that are not better explained by other factors:

- choking or gasping during sleep,
- recurrent awakenings from sleep,

- unrefreshing sleep,
- daytime fatigue,
- impaired concentration; and/or

C. Overnight monitoring demonstrates five or more obstructed breathing events per hour during sleep. These events may include any combination of obstructive apneas/hypopneas or respiratory effort related arousals

3.4. Severity levels of OSA

Severity of the OSA has two components: severity of daytime sleepiness and of overnight monitoring. A severity level should be specified for both components. The rating of severity for the syndrome should be based on the most severe component ("Practice parameters for the use of portable recording in the assessment of obstructive sleep apnea. Standards of Practice Committee of the American Sleep Disorders Association," 1994):

A. Sleepiness

1. Mild: Unwanted sleepiness or involuntary sleep episodes occur during activities that require little attention.

Symptoms produce only minor impairment of social or occupational function.

2. Moderate: Unwanted sleepiness or involuntary sleep episodes occur during activities that require some attention.

Symptoms produce moderate impairment of social or occupational function.

3. Severe: Unwanted sleepiness or involuntary sleep episodes occur during activities that require more active attention.

Symptoms produce marked impairment in social or occupational function.

B. Sleep Related Obstructive Breathing Events

1. Mild: 5 to 15 events per hour

2. Moderate: 15 to 30 events per hour

3. Severe: greater than 30 events per hour

3.5. Portable Monitors

Based in the technology used or the signals that allows achieve, sleep studies are subdivided in 4 levels ("Practice parameters for the use of portable recording in the assessment of obstructive sleep apnea. Standards of Practice Committee of the American Sleep Disorders Association," 1994) (Table 3) determinate through an evaluation of the evidence on portable monitors (PM) as an alternative to in-laboratory polysomnography (PSG) (Collop et al., 2007).

Table 3: Types of sleep studies used in sleep laboratories

Type	Characteristic
1	Full attended polysomnography (≥ 7 channels) in a laboratory setting
2	Full unattended polysomnography (≥ 7 channels)
3	Limited channel devices (usually using 4–7 channels)
4	1 or 2 channels usually using oximetry as 1 of the parameters

Cardiorespiratory monitoring alone can also be used. This involves the measurement of airflow, respiratory effort, oxygen saturation and cardiac frequency, but not electroencephalography (EEG) (Type 3 or 4). The great advantages of these systems are price, portability and the ability of patients to monitor themselves at home (Jennum & Riha, 2009) but they should always be performed only in conjunction with a comprehensive sleep evaluation. Although some studies of portable ambulatory type 3 or 4 monitoring reported a sensitivity of 78–100% and a specificity of 67–100% in comparison with in-lab PSG, ambulatory monitoring is of course not a gold standard. (Ferber et al., 1994) They have some important limitations and so, the PSG, still be considered as the gold- standard and be classified as type 1 of evidence.

Overnight oximetry is sometimes used as a screening test for identifying patients who are at risk of significant OSA, but this should never be seen as a substitute for in-lab PSG or home cardiorespiratory monitoring. There are severe limitations inherent in this technique used in isolation, including the inability to detect apneas or hypopneas not associated with oxygen desaturation and the upper airway resistance syndrome (Jennum & Riha, 2009).

The limitations of the PM are various. They can only be performed in patients with a high pretest probability of moderate to severe OSA without comorbid sleep or medical disorders when all of the previous parameters are met (Collop et al., 2007). In PM type 3 or 4 we don't have access to the sleep stages and, consequently, we can't calculate the apnea-hypopnea index (AHI) because we don't know the total sleep time, and so we calculate the respiratory disturbance index (RDI) (that have a different

definition on the PSG) that considers a number of events, scored as the same rules of the PSG, in the total record time. As result, are likely to underestimate the severity of OSA when compared with PSG because considers that the patient slept in all of the record, what can be false, and don't allow associate the relationship between severity and sleep stage, in particular REM, that exacerbate many times the severity of the apneas and hypopneas (Flemons et al., 2003).

3.6. Risk Factors

3.6.1. Obesity

Epidemiologic studies around the world have consistently identified body weight as the strongest risk factor for obstructive sleep apnea (Al Lawati et al., 2009; Punjabi, 2008). Excess body mass is associated with alterations in upper-airway structure through several distinct mechanisms including: (1) increased parapharyngeal fat deposition resulting in a smaller upper airway, (2) alterations in neural compensatory mechanisms that maintain airway patency, (3) respiratory control system instability, and (4) reduction in functional residual capacity with a resultant decrease in the stabilizing caudal traction on the upper airway (Madani & Madani, 2009; Punjabi, 2008; Stierer & Punjabi, 2005).

Given that, the pathophysiology of obstructive sleep apnea is intimately linked with obesity an estimated 58% of the moderate to severe cases attributable to a Body Mass Index (BMI) greater than/or equal to 30Kg/m² or higher (Doghramji, 2008).

3.6.2. Neck and abdominal circumference

Is frequently described that diameter of the neck and abdominal circumference is an important risk factor for OSA. Obesity / visceral obesity is the major risk factor for the development of OSA, because associated with anatomic alterations that predispose to upper airway obstruction during sleep and reduction in lung volume leads to an increase in pharyngeal collapsibility (Lam, Sharma, & Lam, 2010).

Several studies have shown that the severity of OSA correlates with neck circumference, as a result of excess adipose tissue in the neck compressing the airway. A neck size of > 37 cm for women and 42 cm for men is generally accepted as clinically significant (Davies et al., 1992; Doghramji, 2008).

3.6.3. Gender

It is related that there is a much larger number of men affected by OSA than women. The ratio between men and women in clinical-based studies ranges from 2:1 to 9:1 according to various papers (Jennum &

Riha, 2009). There are several possible reasons for this disparity in sex difference ratio (Madani & Madani, 2009).

Hormonal differences are believed to account for the sex difference in OSA prevalence. Physiologic data suggest that the upper airway in women may be less collapsible than in men, a finding that has been attributed to female sex hormones (Stierer & Punjabi, 2005).

Because sex hormone levels change with menarche, pregnancy and menopause, it is plausible that these changes modify the risk of OSA, most focus on menopause (Kapur, 2010; Lam et al., 2010; Madani & Madani, 2009; Manber & Armitage, 1999; Punjabi, 2008; Stierer & Punjabi, 2005). Pre-menopausal women seems protected from OSA even if they have other known risk factors for OSA (Lam et al., 2010). Furthermore, hormone replacement therapy in post-menopausal women has been associated with a lower prevalence in epidemiologic studies and some studies indicates that the prevalence of OSA is similar to premenopausal women (Lam et al., 2010; Madani & Madani, 2009).

How women and men perceive and relate symptoms of sleep disorders breathing are reported too. If women are less likely to report classic symptoms like loud snoring, nocturnal snorting or gasping, and witnessed apneas, they are less likely to be referred to sleep clinics for evaluation (Punjabi, 2008; Yaggi & Strohl, 2010).

In fact, analyses from different referral centers show that woman with obstructive sleep apnea have a greater tendency to report symptoms of fatigue and lack of energy than men. The different descriptions of OSA symptoms made by female or male bed partner are an important issue too. Female bed partners of male patients appear to have a lower threshold for symptom perception and reporting than male bed partners of female patients (Madani & Madani, 2009; Punjabi, 2008).

There are sex-related differences in upper-airway anatomy and function, obesity and fat distribution and ventilator control. Men have increased fat deposition around the pharyngeal airway as compared with women. Other characteristics like larger tongue size in men's than women's, longer pharyngeal airway in men's contributing to a predispose for an airway obstruction (Kapur, 2010; Madani & Madani, 2009; Punjabi, 2008).

Finally, another hypothesis is the general expectation that the disorder predominantly affects men, takes health care providers to have a lower index of suspicion about the presence of OSA in women's (Punjabi, 2008).

3.6.4. Age

Sleep characteristics changes during life. With advancing age, sleep quality decrease and sleep-related difficulties are more common, often manifested by complaints of difficulty falling asleep, increased number of night-time awakenings, a decreased amount of night-time sleep and lower arousal threshold (Madani & Madani, 2009; Yaggi & Strohl, 2010). Possible explanations for that include age-related changes in upper-airway caliber, attenuation in the ventilatory response to hypoxia and hypercapnia, decreases in functional activity of the upper airway, and an increase in the variability of ventilation during sleep. These changes increase deposition of fat in the parapharyngeal area, lengthening of the soft palate, and changes in body structures surrounding the pharynx, genioglossus negative pressure reflex seems to deteriorate with aging causing airway collapse (Lam et al., 2010; Madani & Madani, 2009; Yaggi & Strohl, 2010).

This age-related variability in sleep stability contributes in some way to an increasing prevalence of OSA with advancing age. However, the existence of OSA in childhood, adolescence and older age means that there is no simple positive correlation of OSA with age (Jennum & Riha, 2009; Yaggi & Strohl, 2010).

The 50-60 age group is the most frequently associated to the presence of OSA. The peak of age is 65 years old (Doghramji, 2008; Kapur, 2010; Madani & Madani, 2009; Yaggi & Strohl, 2010).

3.6.5. Ethnicity

This is a complex factor and several caveats should to be considered in the interpretation of the data linking race with an increased risk for OSA because many things can influence the results. One of the reasons is that populations with a higher prevalence of comorbid medical conditions, including obesity in conjunction with a low socioeconomic status and disadvantages in health care, could explain the higher prevalence of OSA apnea in this groups (Punjabi, 2008).

Other factor can be the differences in craniofacial morphology. Despite lower levels of obesity, Asian samples seem to have a similar prevalence rate compared with the West when we are available ethnic and population differences. Asians have greater disease severity compared with whites. One possible explanation for such differences is in craniofacial morphology that are considered as the etiologic factors for the increased risk and greater severity of OSA in non-obese (Lam et al., 2010; Punjabi, 2008; Stierer & Punjabi, 2005; Yaggi & Strohl, 2010). Asians have soft tissue abnormalities, narrowed nasal cavities and tonsillar hypertrophy play an important role in the development of sleep apnea (Yaggi & Strohl, 2010; T. Young et al., 2004).

3.6.6. Craniofacial and upper-airway structure

As referred above, craniofacial and upper structures are an important issue in the development of sleep apnea (Yaggi & Strohl, 2010; T. Young et al., 2004).

Abnormal characteristics include enlarged tonsils, enlarged uvula, “narrow airway”, macroglossia, retroplaced mandible, soft tissue, tonsillar hypertrophy, inferiorly positioned hyoid bone, and decreased posterior airway space can alter the mechanical properties of the upper airway and increase its propensity to collapse during sleep and promote the occurrence of apneas and hypopneas during sleep (Doghramji, 2008; Lam et al., 2010; Madani & Madani, 2009; Punjabi, 2008; Stierer & Punjabi, 2005; Yaggi & Strohl, 2010; T. Young et al., 2004).

Mallampati scoring consists in a simple noninvasive method used to assess the difficulty of endotracheal intubation is very often used to identify craniofacial and upper-airway structure (Nuckton, Glidden, Browner, & Claman, 2006).

3.6.7. Excessive daytime sleepiness

Sleepiness can be regarded as “normal” sleepiness (a result of the normal circadian rhythm) and “pathological” sleepiness (a result of altered sleep scheduling). Pathological sleepiness can be further subdivided into “habitual” (e.g. as the result of recurring precipitants of sleepiness such as OSA) or “occasional” (e.g. as the result of jet lag or medication) (Jennum & Riha, 2009).

Patients with OSA referred frequently daytime sleepiness in various situations. To evaluate daytime sleepiness in OSA the most widely questionnaire used to assess the likelihood of falling asleep in various situations and the best-validated scale is the Epworth Sleepiness Scale (ESS). The ESS consists in a questionnaire with 8 situations related to daily life and the subject is asked to score (on a 0 to 3 scale) how likely he/she doze off in these situations. There is a maximal score of 24 and normal subjects score between 0 and 10. Scores greater than 10 is usually considered an indication of excessive day time sleepiness, greater than 12 suggest excessive sleepiness and further evaluation of OSA should be considered. The advantages include ease of administration and low cost. It assesses global level of sleepiness and is independent of short-term variations in sleepiness with the time of day and also of inter-day variations. The ESS is also able to discriminate between normal and pathological sleepiness. The accuracy of the ESS depends on the awareness of subjects falling asleep, which may not always be the case (Jennum & Riha, 2009; Johns, 1991).

Magnitude of daytime sleepiness associated with OSA was correlated with crash risk. Untreated OSA is a public risk because increase the risk of traffic accidents and their consequences (Al Lawati et al., 2009; Jennum & Riha, 2009).

3.6.8. Genetics/Family history

The risk factors listed earlier are also “complex” traits, and risk factors can operate either alone or in combination of many, genetic and family history are not an exception.

The genetic influence is multifactorial rather than due to a single mutation or protein action and prohibit definitive conclusions on genetic underpinnings for OSA and that additional studies are needed to further define whether the disorder truly has a genetic component

Some important characteristics such as craniofacial morphology, cephalometric abnormalities, including retroposition of the maxilla and mandible and a large soft palate, volume of the lateral parapharyngeal walls, tongue, soft tissue structures and other factors like self-reported sleepiness, ventilatory control and sleep cycles/ architecture operating during sleep are in part the result of various genetic and environmental factors that act and interact to produce disease (Madani & Madani, 2009; Punjabi, 2008; Stierer & Punjabi, 2005; Yaggi & Strohl, 2010). Although the difficulty to define the genetic basis of obstructive sleep apnea, the available data suggests that inquiries about family history can certainly aid in identifying possible patients due to the familial susceptibility for sleep apnea seems to increase directly with the number of affected relatives (Punjabi, 2008).

3.6.9. Smoking

Airway inflammation and damage due to cigarette smoke could alter the mechanical and neural properties of upper airway and increase its collapsibility during sleep (Lam et al., 2010; Punjabi, 2008; Yaggi & Strohl, 2010). Sleep instability, which has been linked to OSA, may be increased by overnight reductions in nicotine blood levels (Madani & Madani, 2009). Although the association with OSA is relatively weak, smoking may interact with and add to the cardiovascular risk associated with OSA (Yaggi & Strohl, 2010)

3.6.10. Alcohol and sedatives

Alcohol and sedatives ingestion can induce apneic activity in normal or asymptomatic individuals precipitate obstructive apneas and hypopneas during sleep (Punjabi, 2008) because relaxes upper airway dilator muscles and so increases upper airway resistance resulting in hypotonia of the oropharyngeal muscles (Doghramji, 2008).

Therefore, alcohol intake can prolong apnea duration, suppress arousals, increase frequency of occlusive episodes and worsen the severity of hypoxemia (Lam et al., 2010; Madani & Madani, 2009).

3.6.11. Comorbid conditions

Obstructive sleep apnea also has been implicated in the etiology of comorbid and cardiovascular conditions, including hypertension, coronary artery disease, congestive heart failure, and stroke (Punjabi, 2008).

OSA is high in patients with hypertension and a casual role of OSA in hypertension has been suggested in several studies (T. Young et al., 2004). Some studies suggests that there are a potential relationship between OSA and stroke, however, this implication needs to be proved (Punjabi, 2008).

The treatment of OSA can improve the conditions related above, and so, confirm the relationship between these conditions. Some medical conditions such as uncontrolled hypertension, coronary artery disease, congestive heart failure, stroke, and diabetes mellitus, undiagnosed obstructive sleep apnea should be considered as a possible concomitant problem. The reason maybe that intermittent hypoxemia and sleep disruption of obstructive sleep apnea are deleterious to glucose homeostasis and alleviating obstructive breathing during sleep with continuous positive airway pressure therapy has direct effects in improving hyperglycemia and improve the metabolic control (Punjabi, 2008).

3.7. Diagnostic decision support

The definition of clinical decision support systems is now a major topic since it may help the diagnosis, prognosis, and treatment selection. However, the complicated nature of real-world biomedical data has made it necessary to look beyond traditional biostatistics without losing the necessary formality (P. Lucas, 2004). New computational techniques are better at detecting patterns hidden in biomedical data, and can better represent and manipulate uncertainties. For example, naive Bayesian approaches are closely related to logistic regression (Schurink et al., 2007). Bayesian approaches have an extreme importance in these problems as they provide a quantitative perspective, and allow taking into account prior knowledge when analyzing data, offering a general and versatile approach to capturing and reasoning with uncertainty in medicine and health care (P. J. F. Lucas, van der Gaag, & Abu-Hanna, 2004).

3.7.1. Traditional clinical models

When we have a dichotomous outcome, the technique of choice for statistical modeling is logistic regression (LR) (Tu, 1996). The relationship is achieved through the logistic regression equation, which

allows determine which explanatory variables influence the outcome and, consequently evaluate the probability that individual's values of the explanatory variables, will have a particular outcome (Petrie & Sabin, 2009). The usual assumption is that these predictor variables are related in a linear manner to the log odds of the outcome of interest (Tu, 1996).

Given the widely spend use of these models, more details are not present here.

3.7.2. Beyond traditional statistics

Actually, the health care services produce data that is increasing every day as a consequence of new techniques and the integration of data from different sources. The traditional statistics methods, like logistic regression, have been unable to deal with great databases, and so, the utilization of new methods, particularly machine learning ones, has been increasing. Data mining (DM) allows methods for data preprocessing and visualization, non-statistical methods and new methods based on probabilities and statistics that supports clinical decisions on models (P. Lucas, 2004).

Artificial intelligence is a branch of computer science and machine learning is one of its subdivisions and, from the beginning, is used in medical databases. Tom Mitchell defines machine learning as "the study of computer algorithms that improve automatically through experience. Successful applications range from data mining programs that discover general rules from large databases, to information filtering systems that learn users 'reading preferences, to autonomous vehicles that learn to drive on public highways" (Mitchell, 1997). The three main branches of machine learning are statistical and pattern recognition methods like k-neighbors and Bayesian classifiers, inductive learning of symbolic rules like decision trees, decision rules or induction of logic programs and, finally, artificial neural networks (Kononenko, 2001).

The Knowledge Discovery in large Databases process consists of five basic steps: (1) problem identification; (2) data extraction; (3) data preprocessing; (4) data mining, and; (5) pattern interpretation and presentation. The main tasks of data mining in healthcare may include (1) discovering associations, (2) clustering, or (3) creating predictive (classification/regression) models (Lee & Abbott, 2003). DM is concerned with finding patterns in large databases which are interesting and valid. There are numerous data mining algorithms that can be used in classification or prediction- predictive data mining algorithms, or finds associations, clusters- descriptive data mining algorithms (Lavrac, 2001).

Decision support methods are built based in model selection that uses the best algorithm for a given dataset, and model integration/combination. They can provide an optimal solution and can be applied to build rules or decision trees proposing the best classifier for a given classification task.

In the healthcare/medical domain DM tools commonly used include neural networks, decision trees and Bayesian networks. Neural networks are designed to mimic the parallel processing ability of the human brain. Decision trees use a repeating series of branches that describes associations between attributes and a target variable. Bayesian networks provides a probabilistic approach (Lee & Abbott, 2003).

To be applied in medical diagnostic tasks, machine learning based systems have to some requirements (Kononenko, 2001):

- 1) Good performance: high values for diagnostic accuracy is crucial. The performance of most algorithms is at least equal to physicians.
- 2) Dealing with missing data: In some patients records have lack of information, so, dealing with incomplete descriptions of patients is important.
- 3) Dealing with noisy data: Uncertainty and errors are common in medical data. Therefore, ML applications must have effective means for handling noisy data.
- 4) Transparency of diagnostic knowledge: the problem can be presented by the system in a different and new point of view, transparent to physician, not see before in an explicit form.
- 5) Explanation ability: diagnosis must be presented in a clear way when diagnosis new patients.
- 6) Reduction the number of tests: patient history has a large amount of data. The classifier must be able to reliably diagnose with a small amount of data about the patients.

3.8. Bayesian Networks

Bayesian networks (BN) are graph-based formalisms for the representation and manipulation of uncertain knowledge, based on probability theory. They provide a probabilistic approach to inference which allow taking into account prior knowledge when analysing data (P. Lucas, 2004; Mitchell, 1997). They are important to machine learning because they provide a quantitative approach to weighing the evidence supporting alternative hypothesis and represent a joint probability distribution and domain (or expert) knowledge in a compact way (Lee & Abbott, 2003; Mitchell, 1997). They consist of a qualitative and quantitative part. The qualitative part encodes, in a directed graph, the variables under study with their probabilistic interrelationships. The quantitative part is a set of conditional probabilities describing the strengths of the dependences between variables represented in the qualitative part. This parts together are sufficient to define a joint probability distribution on the statistical variables under study (Coupé & van der Gaag, 2002).

BN are one of the most popular uncertainty formalisms because: they can handle noise, missing information and reveal probabilistic relations; possibilities learn from data and incorporate domain knowledge and provide a good interface through their compact graphical representation. Another advantage is that networks are flexible and the learned models can be used for many tasks like prediction

or diagnosis. One practical difficulty in applying Bayesian methods is that they typically require initial knowledge of many probabilities. A second practical difficulty is the significant computational cost required to determine the Bayes optimal hypothesis in the general case (Mitchell, 1997).

For all the characteristics referred above and also that they are a powerful data mining technique for handling uncertainty in complex domains and a fundamental technique for pattern recognition and classification, the use of BN are increasing in the choice of data mining techniques applied in medical domain (Lee & Abbott, 2003). They allow the stepwise combination of prognostic evidence and provide a quantitative measure in terms of probabilities (Sakellaropoulos & Nikiforidis, 2000).

3.8.1. Probabilistic reasoning

To understand BN some basic probabilistic concepts have to be learned.

If X and Y are random variables, with probability distributions $P(X)$ and $P(Y)$, the joint probability distribution $P(X, Y)$ represents the distribution of both variables related. This way, for a given event $X = x$, the marginal probability of $X = x$, can be calculated:

$$P(X = x) = \sum_Y P(X = x, Y).$$

The conditional probability is a concept very important in medicine and is defined as an event $X = x$ given the occurrence of $Y = y$:

$$P(X = x | Y = y) = \frac{P(X, Y)}{P(Y)}.$$

Given the properties of the joint distribution, the former equation can be rewritten as a famous theorem.

The Bayes theorem provides a direct method for calculating the posterior probabilities of the various hypotheses given observations. More precisely, Bayes theorem provides a way to calculate the probability of a hypothesis $P(h | D)$ based on its prior probability $P(h)$, the probability of observing various data given the hypothesis $P(D | h)$, and the observed data itself (Mitchell, 1997), making it the cornerstone of Bayesian learning methods:

$$P(h | D) = \frac{P(D | h)P(h)}{P(D)}.$$

In many scenarios, the learner is interested in finding the most probable hypotheses h of a set of candidate hypotheses H given the observed data. The maximally probable hypothesis is called *maximum a posteriori* (MAP) hypothesis:

$$h_{MAP} = \arg \max_{h \in H} P(D | h)P(h).$$

In other cases, we assume that every hypothesis in H is equally probable a priori. So, we just need consider the term $P(D | h)$, called the *likelihood* of the data D given h , to find the most probable hypothesis, and any hypothesis that maximizes $P(D | h)$ is called a *maximum likelihood* (ML) hypothesis, h_{ML} :

$$h_{ML} = \arg \max_{h \in H} P(D | h).$$

The most probable classification of the new instance is obtained by combining the predictions of all hypotheses, weighted by their posterior probabilities. The Bayes optimal classification of the new instances v_j is described by (Mitchell, 1997):

$$\arg \max_{v_j \in V} \sum P(v_j | h_i)P(h_i | D).$$

The *Bayes optimal classifier* or Bayes optimal learner is every system that classifies new instances based in the equation above.

When multiple dependences are stake, conditional independence is defined as a logic that supports symbolic reasoning about dependence and independence information, making it possible to abstract away from the numerical detail of probability distributions and the process of assessing probability distributions.

Let X, Y, Z be sets of variables, X is conditionally independent of Y given Z if:

$$P(X | Y, Z) = P(X | Z).$$

This feature is the cornerstone of BN learning.

3.8.2. BN: Definition, representation

Bayesian network is a graphical representation of stochastic (statistical) dependences and independences among variables. The type of dependence and independence we are dealing with is determined by the direction of arcs and whether or not particular variables are instantiated. It is based on the assumption that the classification of patterns is expressed in probabilistic terms between predictors and outcome variables-

conditional assumption (Lee & Abbott, 2003). Clearly, the assumptions that are made in formulating this prior knowledge are crucial, and are a topic of much debate (P. Lucas, 2004).

On the BN representation, graphical information is qualitative, nodes represents variables and arcs specify the (in) dependence between variables (Sakellariopoulos & Nikiforidis, 2000). In this way, each node is independent of all its non-descendent nodes, given its parents. This causes the joint probability distribution P is equivalent to the product of the (conditional) distributions (Sakellariopoulos & Nikiforidis, 2000):

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i | \pi_{x_i}).$$

Where π_{x_i} is the set of parents of the vortex corresponding to the variable X_i .

So, A Bayesian *network* \mathbf{B} is defined as a pair $\mathbf{B} = (G, P)$, where $G = (V(G), A(G))$ is an acyclic directed graph with a set of vertices (or nodes) $V(G) = \{X_1, X_2, \dots, X_n\}$ and a set of arcs $A(G) \subseteq V(G) \times V(G)$, and where P is a joint probability distribution defined on the variables corresponding to the vertices $V(G)$. The basic property of a Bayesian network is that the joint distribution $P(X_1, X_2, \dots, X_n)$ is equivalent to the product of the (conditional) probabilities:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \pi_{X_i}).$$

Thus, $P(X_i | \pi_{X_i})$ are the (conditional) probability distributions witch are specified for the variable X_i , for $i = 1, \dots, n$, in creating a Bayesian network (Coupé & van der Gaag, 2002).

3.8.3. Building BN

The construction of a BN has two phases (Lee & Abbott, 2003):

1. Creation of a BN structure: a acyclic graph which encodes probabilistic relationships among variables;
2. Assessment of the prior and local conditional probabilities: training and testing the network structure.

In contrast with logistic regression, where dependence and independence is hidden in approximating weights, in BN structure these are explicitly represented (Lee & Abbott, 2003).

We can construct a BN manually or learn from data. Manual construction in practice are time consuming because requires access of human experts. Learning from data in nowadays are much more attractive, consequence of the increasing of clinical and biological data (P. J. F. Lucas et al., 2004).

The manual way comprehends various stages using the expert knowledge, relevant literature and analysis of available patient data (P. J. F. Lucas et al., 2004). The five main stages are:

1. Selection of relevant variables: is generally based on interviews with experts, descriptions of the domain and an extensive analysis of the purpose of the network under construction.
2. Identification of the relationships among the variables: determine how those factors are related to each other. Dependence and independence relationships between them have to be analysed and expressed in a graphical structure. Causal graph, common effects and causes.
3. Qualitative probabilistic and logical constrains: qualitative probabilistic derived from properties of stochastic dominance of distributions. Logical constrains are derived from functional relationship between the variables.
4. Assessment of probabilities: Local conditional probability distributions $Pr(X_i | \pi(X_i))$ for each variable X_i are filled in. The required probabilities can be obtained from domain experts or, alternatively, from data.
5. Sensitivity analysis and evaluation: to be used in real-life practice, BN are tested and evaluated. One way to assess network's quality is to perform a sensitivity analysis with patient data. There are various ways to evaluate BN like measuring classification performance on a given set of real patient data and measuring similarity of structure or probability distribution to a gold-standard network or other probabilistic model.

3.8.4. Learning BN from data

BN can be learnt from data without explicit access to knowledge of human experts by exploring various issues such as comparison of learning algorithms, dealing with missing data and evaluation of the networks learned (P. J. F. Lucas et al., 2004).

To create BN from data with learning purpose this have to satisfy some requisites: data collection is very important to avoid bias that interfere in the BN impact and purpose; data's variables and values should match characteristics to be modelled in the network or should at least admit easy translation; the size of the data have to allow reliable information of probabilistic relationships among variables discerned; must have properties that allows the use of the most learning algorithms.

One important aspect is that many statistical and learning methods cannot deal with missing values and the absence of missing values has to be ensured by two ways: removing the cases with missing data or

filling (imputing) missing data. The first method has to be applied with caution as it can result in the loss of a large amount of valuable data, thus leading to a decrease in the robustness of the models learned. The second one means that the missing value is replaced with an estimate of the actual value (P. J. F. Lucas et al., 2004).

Learning Bayesian networks involves both structure learning, i.e., learning the graph topology from data, and parameter learning, i.e., learning the actual, local probability distributions from data. There are basically two approaches to structure learning: search and score structure learning, and constraint-based structure learning.

Search-and-score algorithms search for a BN structure that fits the data best (in some sense). These methods search the space of all possible acyclic digraphs by generating various different graphs in a heuristic way and comparing these to their ability to explain that at hand. They start with an initial network structure (often a graph without arcs or a complete graph), and then traverse the search space of network structures by in each step removing an arc, adding an arc, or reversing an arc. Recent search-and-score-algorithms take Markov equivalence into account, i.e., they search in the space of equivalence classes of Bayesian networks and the scoring method they use give the same score for equivalent networks. Bayesian networks with different graph topologies that are included in the same Markov equivalence class represent exactly the same conditional-independence information by d-separation. Examples of search and score algorithms are K_2 and inclusion-driven learning. They usually are based on hill climbing (greedy) search (P. J. F. Lucas et al., 2004). K_2 performs a greedy search that trades off network complexity for accuracy over the training data (Mitchell, 1997).

Constraint-based algorithms carry out a conditional (in) dependence analysis on the data and allow for the easy incorporation of background knowledge, i.e., prior knowledge on dependences or independences that hold for the domain under consideration. Examples of constraint-based learning algorithms are PC, NPC, growshrink, and incremental association (P. J. F. Lucas et al., 2004).

3.8.5. Naïve Bayes classifier

The naïve Bayes (NB) classifier is based on the simplifying assumption that the attribute values are conditionally independent given the target value (Mitchell, 1997):

$$v_{NB} = \arg \max_{v_j \in V} P(v_j) \prod_i P(a_i | v_i)$$

One interesting difference between the naïve Bayes learning methods and other learning methods is that there is no explicit search through the space of possible hypothesis (Mitchell, 1997). Naïve Bayes is robust to the presence of irrelevant attributes (but redundant variables must be taken into account, as they have impact on performance), the variability of a data set is summarized in contingency tables and the dimension of the decision model is independent of the number of examples.

The general structure of a naïve Bayesian network is shown of figure 1.

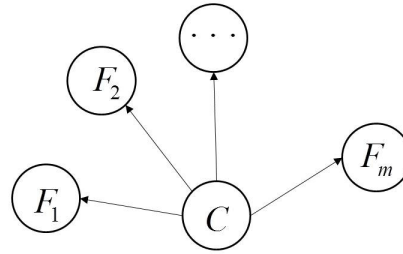


Figure 1: Naive Bayes networks

F represents the features variables and C the class variable

3.8.6. Tree augmented Bayesian network

Tree augmented Bayesian network (TAN) is an extension of naïve Bayes: reducing the number of independent assumptions, each node has at most two dependences, one conditionally from the class and other conditionally from other attribute (P. J. F. Lucas et al., 2004) (figure 2).

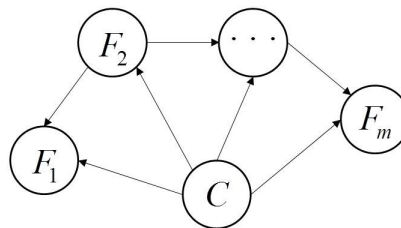


Figure 2: Tree augmented Bayesian network

F represents the features variables and C the class variable

TAN computes the Mutual Information between all pairs of variables given the class to define the network structure:

$$I(X, Y | C) = \sum_c P(c) I(X, Y | C = c).$$

3.8.7. Probabilistic inference

If a new case is inserted in the BN, new evidence is introduced to the BN which causes an update of the belief in various prognostic outcomes. System calculates the posterior probability of the outcomes, given the specific evidence. This process is called inference and BN are used with this function (Sakellariopoulos & Nikiforidis, 2000). It means that we can infer the value of some target variable given the observed values of other variables using BN.

In general, a BN can be used to compute the probability distribution for any subset of network variables given the values or distributions for any subset of the remaining variables. What we need to infer is the probability distribution for the target variable, which specifies the probability that it will take on each of its possible values given the observed values of the other variables (Mitchell, 1997).

3.9. Performance measures in classification problems

To improve performance, determine an appropriate sensitivity-specificity trade-off, through applying different algorithms and parameters is crucial (Lavraç, 2001). The performance and quality of predictive models should be evaluated by their abilities of discrimination and calibration (Dreiseitl & Ohno-Machado, 2002; Lee & Abbott, 2003).

Discrimination measures how much the model is able to separate cases with positive outcome value from those with negative outcome value. Calibration is a measure of how close the predicted values are to the real outcomes, measuring whether they are high or low when compared to the real outcomes (Dreiseitl & Ohno-Machado, 2002; Lee & Abbott, 2003).

The most commonly used measures of discriminatory power are the area under the ROC curve (AUC), a plot of the sensitivity versus specificity of a model in a binary classification task sensitivity, specificity, and accuracy (Dreiseitl & Ohno-Machado, 2002; Lee & Abbott, 2003). Calibration of the models can be measured by construction of calibration curve or computation of the Hosmer–Lemeshow goodness-of-fit χ^2 statistic (Lee & Abbott, 2003). Sensitivity is defined as the number of correctly classified cases as positives divided by the total number of actual positive cases. Specificity is defined as the number of

correctly classified cases as negatives divided by the total number of actual negative cases (Lee & Abbott, 2003).

As each sensitivity and specificity are reported for a single threshold point, the ROC curve can be plotted through various cut-off values. AUC gives a definitive measure of the classifiers discrimination ability that is not dependent on the choice of cutoff point value and so represents a measure of sensitivity and specificity over all possible thresholds (Dreiseitl & Ohno-Machado, 2002; Lee & Abbott, 2003).

On health domain the false alarm rate, specificity, needs to be minimized, and the detection rate, sensitivity, which needs to be maximized (Lavrac, 2001).

Accuracy, calculated using a threshold closest to (0, 1) that minimizes the sum of $(1 - \text{sensitivity})^2$ and $(1 - \text{specificity})^2$, is the only discrimination measure influenced by the class distribution in the data set. Aware when the case distribution in the training set is different from the case distribution of the population on which the classifier is used (Dreiseitl & Ohno-Machado, 2002; Lee & Abbott, 2003).

Positive predictive value (PPV) and negative predictive value (NPV) cannot be ignored in reporting the results. PPV is the proportion of cases that the network classifies as positive that actually are positive, and NPV is the proportion of cases that the net classifies as negative that are actually negative (Lee & Abbott, 2003).

To provide an unbiased estimate of a models discrimination and calibration, these values have to be calculated from a data set not used in the model building process. Usually, the original data set is split into test or validation set, and training set to avoid overfitting. There are many ways to evaluate models performance; we will numerate only cross-validation and bootstrapping. On the cross-validation process n-fold cross-validation builds n models; the numbers reported are the averages over all n test sets, learning is done on subsequent parts of the data not used in the train process. The extreme case of using only one data item for testing is known as leave-one-out cross-validation. An alternative to cross-validation is bootstrapping, a process by which training sets are sampled with replacement from the original data sets (Dreiseitl & Ohno-Machado, 2002; Lavrac, 2001).

4. Material and Methods

A literature review was performed to define the relevant variables to be collected. Data was prospectively collected from patients in two periods of time. A logistic regression (LR) and Bayesian networks (BN) models were built. Results from polysomnography (PSG) were categorized into normal or obstructive sleep apnea (OSA) (independently of the severity) and results were validated to improve the generalization of the model using cross-validation (CV). Finally, we apply the final models, LR and BN, on the second comparable cohort.

4.1. Variables

The selection of variables were based in a literature search on Pubmed on 27th of September of 2011, using Mesh terms “risk factors”, “obstructive sleep apnea” and “diagnosis”. We obtained 442 articles and we selected only reviews to determine the variables to our study. At the first search we found 127 reviews and excluded 96 articles after read the title and abstract because they studied children, pregnant, OSA as a secondary problem from another disease, only about women or elderly, or were not related to risk factors. Of the 48 full text required, we excluded 17 because 6 we didn’t had access in time, the others weren’t about risk factors, language not Portuguese or English or were about OSA as a resultant disease. From the bibliography of these articles we included one article, finishing our search with 32 articles.

The 33 collected and studied variables were considered from 32 articles and are presented in table 4 (Berry, 2008; Bonekat & Hardin, 2003; Doghramji, 2008; Epstein et al., 2009; Fogel & White, 2000; Jennum & Riha, 2009; Kapur, 2010; Krieger & Redeker, 2002; Krug, 1999; Kuhlmann, Bormann, & Becker, 2009; Lam & Ip, 2010; Lam et al., 2010; Madani & Madani, 2009; Olson, Park, & Morgenthaler, 2005; Pagel, 2008; Palmer & Redline, 2003; Park, Ramar, & Olson, 2011; Pillar & Lavie, 2011; Punjabi, 2008; Schwab, 2005; Sheldon, Belan, Neill, & Rowland, 2009; Stierer & Punjabi, 2005; Tasali, Mokhlesi, & Van Cauter, 2008; Tate & Tasota, 2002; Togeiro et al., 2010; Yaggi & Strohl, 2010; Younes, 2003; T. Young et al., 2004).

Table 4: Variables collected in the studied patients

Demographic	Physical Examination
Gender	Body Mass Index (BMI)
Ethnicity	Neck circumference (NC)
Age	Abdominal circumference (AC)
	Craniofacial and upper airway abnormalities
Clinical history	Comorbidities and other conditions
Snore	Atrial fibrillation
Witnessed apneas	Stroke
Gasping/Shocking	Myocardial infarction
Motor Vehicle Crashes	Pulmonary hypertension
Refreshing Sleep	Congestive heart failure
Humor alterations	Diabetes
Nocturia	Metabolic Syndrome
Restless Sleep	Renal failure
Decreased libido	Hypothyroidism
Morning Headaches	Gastroesophageal reflux
Alcohol before sleep	Hypertension
Smoker	
Sedative use	
Epworth Somnolence Scale (ESS)	
Concentration decrease	

4.2. Patients

This study included the patients referred to perform PSG at the Sleep Laboratory of Vila Nova de Gaia/Espinho Hospital Center, Portugal. We collected two samples with patients that agreed to participate in the study, one for the development of the models and a second one to test the proposed models. The first sample includes patients that realized PSG between December of 2011 to February of 2012 and the second between April and May of 2012. All adults, older than 18 years, referred by the physicians with suspected OSA and agreed to participate in the study, were included. In case of duplicate studies (sometimes the patient repeats the test), the one with best sleep efficiency was selected. Patients with suspicion of another disorder than OSA, patients already diagnosed (therapeutic studies), and patients with severe lung disease or neurological condition that somehow affects the respiratory function, such as neuromuscular diseases, were excluded.

This study was approved by the Ethics Commission of Centro Hospitalar de Vila Nova de Gaia/Espinho, Portugal and patients were not exposed to any adverse effects.

4.3. Data collection

Clinical information was collected prospectively during consultation, 3 months before PSG. However, all variables were checked to avoid missing information and if some variable was missing it would be collected during PSG. During PSG the parameters, settings, filters, technical specifications, sleep stage, event scoring and final results were applied according to the American Academy of Sleep Medicine (AASM) rules (Iber C, 2007).

To achieve cerebral activity, we used a combination of parameters including EEG, electro-oculography (EOG) and electromyography (EMG). The EEG derivations were C4-M1, O2-M1, C3-M2 and O1-M2. A thermal flow was used to score apneas, a nasal air pressure device to identify hypopneas, a snore sensor for snoring, one derivation for electrocardiography (ECG), chest and abdominal belts to evaluate respiratory effort, position sensor, leg EMG and pulse oximetry.

Hypopneas were scored if the nasal pressure signal excursions drop by $\geq 30\%$ of baseline and duration of this drop was at least 10 seconds with a $\geq 4\%$ desaturation from pre-event baseline (recommended rules of AASM).

Data quality was checked and validated before analysis to avoid duplicate records and errors on data.

For this study, the outcome measure is the clinical diagnosis supported on PSG results, categorized into normal or OSA (mild, moderate and severe).

4.4. Logistic Regression model

Odds ratios (OR) and respective 95% confidence intervals were calculated using univariate LR for the 33 variables as independent variables and the presence of OSA as dependent variable. The independent variables with significant OR were used in the multiple forward conditional logistic regression analysis to construct the model. A significance level of 5% was used.

Receiver Operating Characteristic (ROC) curve analysis was performed to determine prediction error and AUC to achieve the cutoff for our model. The target sensitivity is 100%, to avoid false negatives, yet aiming at reducing the number of unnecessary exams (increasing specificity), but we also tested our results aiming a 95% of sensitivity.

4.5. Bayesian Network model

Models were built using only the significant variables achieved on the univariate LR. First, we made a pre-processing of the data. Missing data was removed and continuous variables Body Mass Index (BMI), Neck circumference (NC) and Abdominal circumference (AC) were categorized into dichotomous variables: BMI into normal or obese, NC and AC into normal or increased, according to the literature (Direcção Geral da Saúde, 2012; Davies et al., 1992) (table 5).

Table 5: Parameters used to recode continuous variables BMI, NC and AC into dichotomous

	Male	Female
Body Mass Index (BMI)		
<30	Normal	Normal
≥ 30	Obese	Obese
Neck Circumference (NC)		
≤ 37	Normal	Normal
37-42	Normal	Increased
>42	Increased	Increased
Abdominal Circumference (AC)		
≤ 80	Normal	Normal
80-94	Normal	Increased
≥ 94	Increased	Increased

To see the relationship among the six variables we build a network without any class information.

The result of PSG (normal or OSA) was defined as the label attribute to construct the models.

The Bayesian networks (BN) classifiers used to build the models were Naïve Bayes (NB) and Tree Augmented Bayesian Network (TAN), given their previous good results in other clinical domains (E. Burnside et al., 2000; Montironi et al., 1995; Sakellaropoulos & Nikiforidis, 1999). Different cutoffs were used to achieve two levels of sensitivity: 100% and 95%.

We use *R* to learn a network without labelling the examples, using the package *BNLEARN*, which provides a free implementation of some BN structure learning algorithms (Scutari, 2010). *Rapidminer* was used to build and evaluate NB and TAN performance and check conditional probabilities given the class outcome of PSG (normal or OSA). Finally, *SamIam* was used to inspect and consult the BNs.

4.6. Models validation

We use the results of sensitivity and specificity to determine the performance of our models and choose different thresholds, accordingly.

The LR model was evaluated with sensitivity and specificity estimates on train data, and using 10 times 2-fold cross validation to check for external validation. All the analysis was performed with SPSS statistical software (SPSS, Inc, Chicago, IL, USA). To evaluate the BN performance models we used leave-one-out CV.

An external validation was made applying the final models, LR and BN, on a second comparable cohort.

5. Results

5.1. Patients

On the first data collection, used to create the logistic regression (LR) and Bayesian networks (BN) models, from the 113 patients considered for inclusion, 27 were excluded for several reasons depicted in figure 3. We collected data from 86 patients, 69 (80%) of which were male and mean age was 56 years. Forty one patients (48%) had normal result with age mean of 54 years; of the 45 patients with obstructive sleep apnea (OSA) (52%), 17 (37%) were categorized into mild, 15 (33%) were moderate and 13 (30%) were severe, and the mean age was 57 years (table 6).

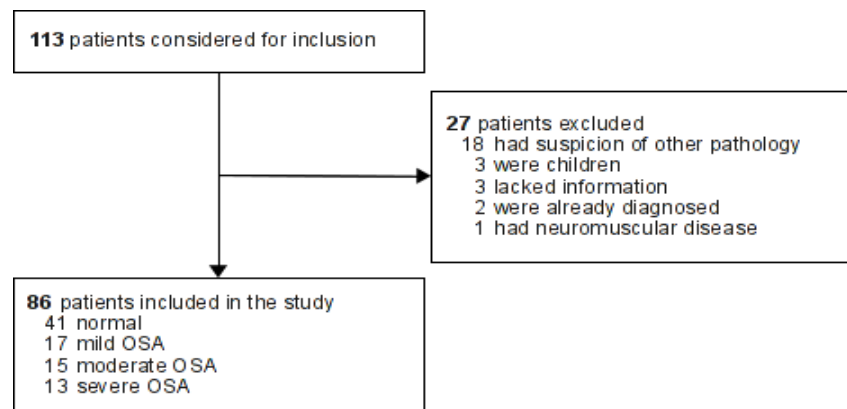


Figure 3: Flow diagram for inclusion of patients in the study

The analysis of univariate regression showed 6 variables (table 6) with significant odds ratio (OR): male gender (OR=7.259, 95% CI=[1.096; 27.651]), body mass index (OR=1.159, [1.030; 1.303]), neck circumference (OR=1.341, [1.159; 1.550]), abdominal circumference (OR=1.076, [1.025; 1.129]), witnessed apneas (OR=4.725, [1.772; 12.599]) and alcohol before sleep (OR=3.307, [1.350; 8.100]).

Table 6: Description and odds ratios for the 33 studied variables

	Normal (n=41)	OSA (n=45)	Simple OR	95%CI
Gender, <i>n (%)</i>				
Male	27 (66)	42 (93)	<i>Ref.</i>	
Female	14 (34)	3 (6)	7.259	[1.096;27.651]
Ethnicity, <i>n (%)</i>				
European	40 (98)	44 (98)	-	-
African	1 (2)	1 (2)	-	-
Age, <i>mean (sd)</i>	54 (14)	57 (13)	1.020	[0.988;1.052]
Snore, <i>n (%)</i>	41 (100)	45 (100)	-	-
Witnessed apneas, <i>n (%)</i>	20 (49)	36 (82)	4.725	[1.772;12.599]
Gasping/Shocking, <i>n (%)</i>	7 (17)	14 (31)	2.194	[0.783;6.142]
Motor Vehicle Crashes, <i>n (%)</i>	3 (8)	3 (7)	0.872	[0.165;4.608]
Refreshing Sleep, <i>n (%)</i>	17 (41)	22 (49)	1.350	[0.575;3.169]
Humor alterations, <i>n (%)</i>	2 (5)	2 (4)	0.907	[0.122;6.751]
Nocturia, <i>n (%)</i>	16 (39)	16 (36)	0.862	[0.359;2.069]
Restless Sleep, <i>n (%)</i>	4 (10)	9 (20)	2.312	[0.653;8.185]
Decreased libido, <i>n (%)</i>	0 (0)	1 (2)	-	-
Morning headaches, <i>n (%)</i>	10 (24)	8 (18)	0.670	[0.236;1.906]
Alcohol before sleep, <i>n (%)</i>	12 (29)	26 (58)	3.307	[1.350;8.100]
Smoker, <i>n (%)</i>				
No	22 (54)	25 (56)	<i>Ref.</i>	
Yes	10 (24)	7 (16)	0.616	[0.200;1.894]
Ex-smoker	9 (22)	13 (29)	1.27	[0.456;3.543]
Sedative use, <i>n (%)</i>	8 (18)	9 (20)	1.031	[0.356;2.986]
ESS, <i>median (range)</i>	8 (19)	8 (24)	0.980	[0.908;1.050]
Concentration decrease, <i>n (%)</i>	8 (19)	3 (7)	0.295	[0.072;1.198]
BMI, <i>mean (sd)</i>	28 (4)	30 (5)	1.159	[1.030;1.303]
NC, <i>mean (sd)</i>	39 (3.4)	43 (3.8)	1.341	[1.159;1.550]
AC, <i>mean (sd)</i>	100 (10)	108 (12)	1.076	[1.025;1.129]
Craniofacial and upper airway abnormalities, <i>n (%)</i>	17 (41)	28 (62)	2.325	[0.979;5.526]
Atrial fibrillation, <i>n (%)</i>	1 (2)	1 (2)	0.909	[0.055;15.020]
Stroke, <i>n (%)</i>	1 (2)	2 (4)	1.860	[0.162;21.319]
Myocardial infarction, <i>n (%)</i>	4 (9)	2 (4)	0.430	[0.075;2.484]
Pulmonary hypertension, <i>n (%)</i>	0 (0)	0 (0)	-	-
Congestive heart failure, <i>n (%)</i>	1 (2)	0 (0)	-	-
Diabetes, <i>n (%)</i>	8 (19)	9 (20)	1.031	[0.356;2.986]
Metabolic Syndrome, <i>n (%)</i>	0 (0)	0 (0)	-	-
Renal failure, <i>n (%)</i>	0	0	-	-
Hypothyroidism, <i>n (%)</i>	3 (7)	0	-	-
Gastroesophageal reflux disease, <i>n (%)</i>	2 (5)	3 (7)	1.393	[0.221;8.783]
Hypertension, <i>n (%)</i>	21 (51)	22 (49)	0.911	[0.391;2.124]

BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference; ESS: Epworth Somnolence Scale; OR: Odds Ratio; CI: Confidence Interval

5.2. Logistic Regression Model

A multiple forward conditional logistic regression analysis was created with neck circumference (NC), gender, witnessed apneas (WA) and consume of alcohol before sleep. Abdominal circumference (AC) and body mass index (BMI) variables were not considered for this model given their high co-linearity with the strongest variable, NC. After two steps, NC and WA were the final variables present in the equation of the multivariate regression, with intercept -11.147 and coefficients 0.256 (OR=1.292) and 1.134 (OR=3.108), respectively.

The ROC curve (fig. 4) analysis demonstrated an AUC of 80%, with a confidence interval (CI) of [70%; 89%]. Given the good discriminative power of the model, a cutoff value of 10% (patients with probability of OSA higher than 10% were recommended PSG) was chosen to achieve a sensitivity of 100% [92%; 100%] and a specificity of 5% [1%; 15%]. However, aiming at a not so strict value for sensitivity, we could get better results for specificity. Actually, with a cutoff value of 25%, a sensitivity of 95% [86%; 99%] and specificity of 35% [22%; 50%] were achieved. Inspecting erroneous classifications with the cutoff value of 25%, the two misclassified OSA patients were actually diagnosed with mild OSA (representing 12% of total mild OSA patients). For the 10% cutoff value the cross-validation results were $98\pm3\%$ sensitivity and $11\pm3.5\%$ specificity, while for the 25% cutoff value the resulting sensitivity was $89\pm4\%$ and $34\pm7\%$ for specificity.

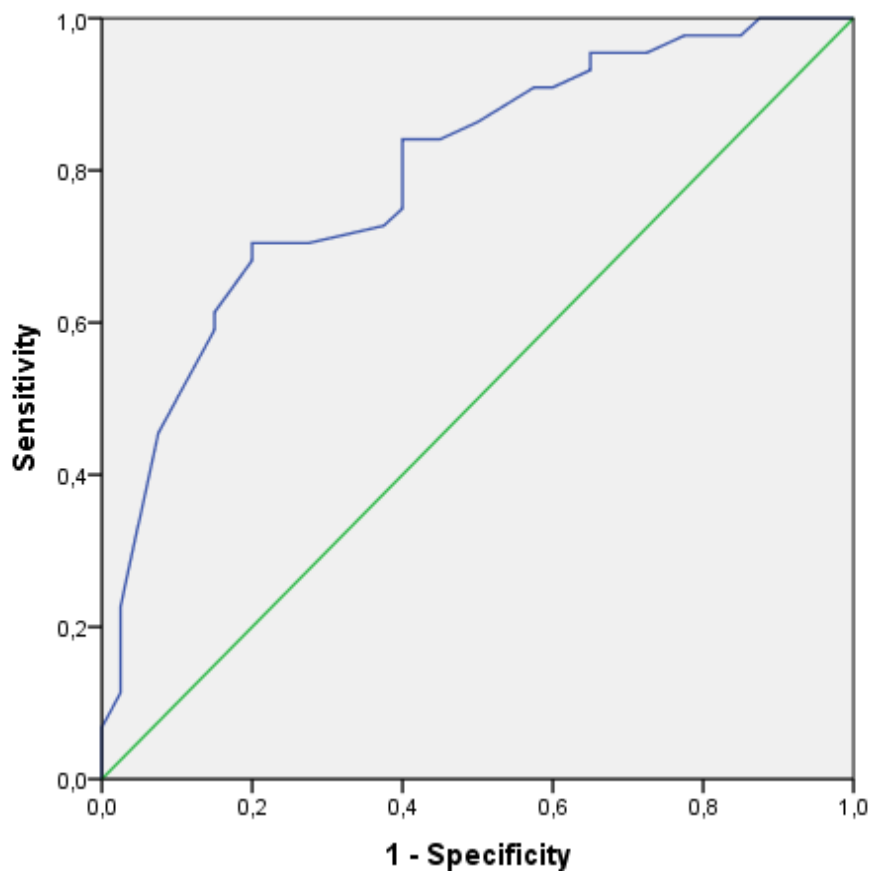


Figure 4: ROC curve of the model

5.3. Bayesian Networks

Figure 5 represents the Bayesian network and the conditional probabilities of the 6 significant variables with significant OR achieved on univariate LR, male gender, NC; AC, WA and alcohol before sleep, created with no class information.

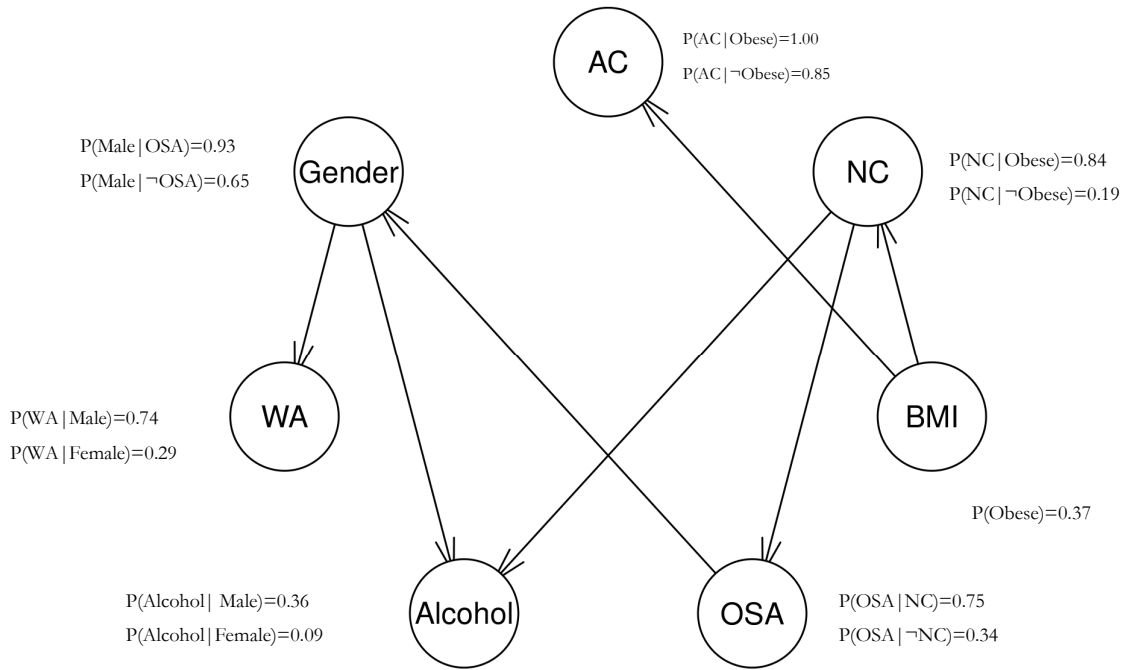


Figure 5: Bayesian network and the conditional probabilities

(WA: Witnessed Apneas, BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference)

As we found on multiple LR, there is a high association between AC, BMI and NC. BMI influences NC and AC, with certain that an obese have AC increased and a higher probability of have a NC increased, $P(\text{NC} | \text{Obese}) = 0.84$, when compared with not obese, $P(\text{AC} | \neg \text{Obese}) = 0.85$ and $P(\text{NC} | \neg \text{Obese}) = 0.19$.

NC influences OSA and alcohol before sleep. There is a high probability that a patient has OSA given having NC increased, $P(\text{OSA} | \text{NC}) = 0.75$, when compared with an NC normal, $P(\text{OSA} | \neg \text{NC}) = 0.34$.

OSA influences gender, with a high probability of a male having OSA diagnosis, $P(\text{Male} | \text{OSA}) = 0.93$ and lower to be normal, $P(\text{Male} | \neg \text{OSA}) = 0.65$. Of course caution is advised in interpreting such dependences as causation.

Gender influences WA and alcohol before sleep, with the presence of these two variables more likely in men, $P(\text{WA} | \text{Male}) = 0.74$ and $P(\text{Alcohol} | \text{Male}) = 0.36$ than in women $P(\text{WA} | \text{Female}) = 0.29$ and $P(\text{Alcohol} | \text{Female}) = 0.09$.

5.3.1. Naïve Bayes network classifier

Figure 6 shows the NB based model with the outcome of PSG as label attribute.

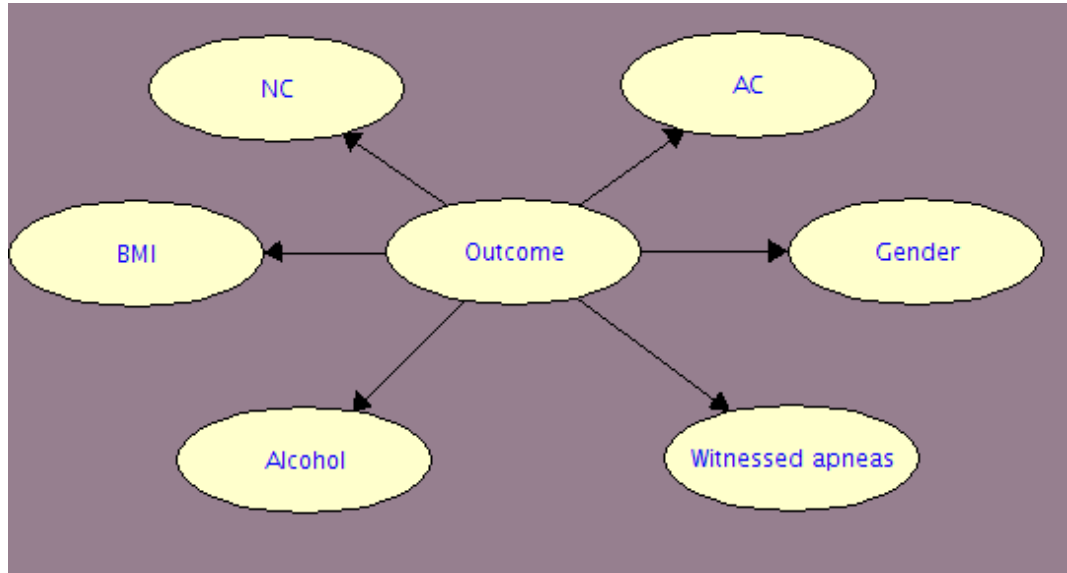


Figure 6: Naive Bayes classifier

(BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference)

Aiming 100% of sensitivity, we achieve 7% as cutoff for NB, obtaining a sensitivity of 100% and 25% for specificity while using a higher cutoff of 10%, for 95% of sensitivity, the results were 98% and 33% for sensitivity and specificity respectively, for internal validation (table 8).

The same cutoffs were used on leave-one-out cross-validation (CV) to check the external validation of NB. Using 7% as cutoff, the results were 98% for sensitivity and 18% for specificity, while using the 10% cutoff the results were 93% to sensitivity and 30% for specificity (table 8).

The marginal probabilities if no information is given to the network are shown on figure 7.

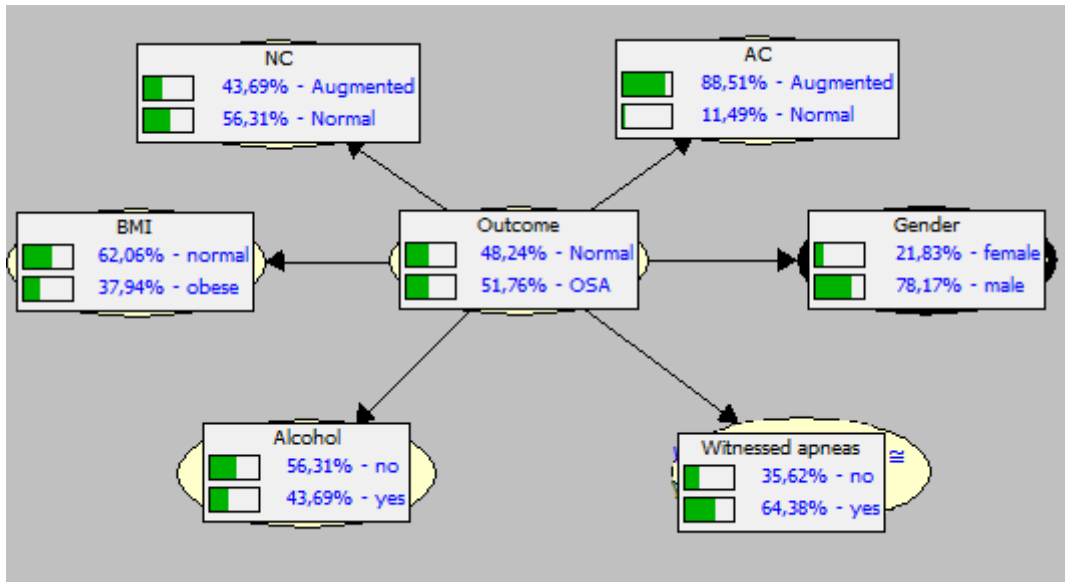


Figure 7: Marginal probabilities of Naïve Bayes classifier to the presence of OSA

(BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference)

5.3.2. Tree Augmented Bayesian network

The tree augmented Bayesian network (TAN) model is shown in figure 8.

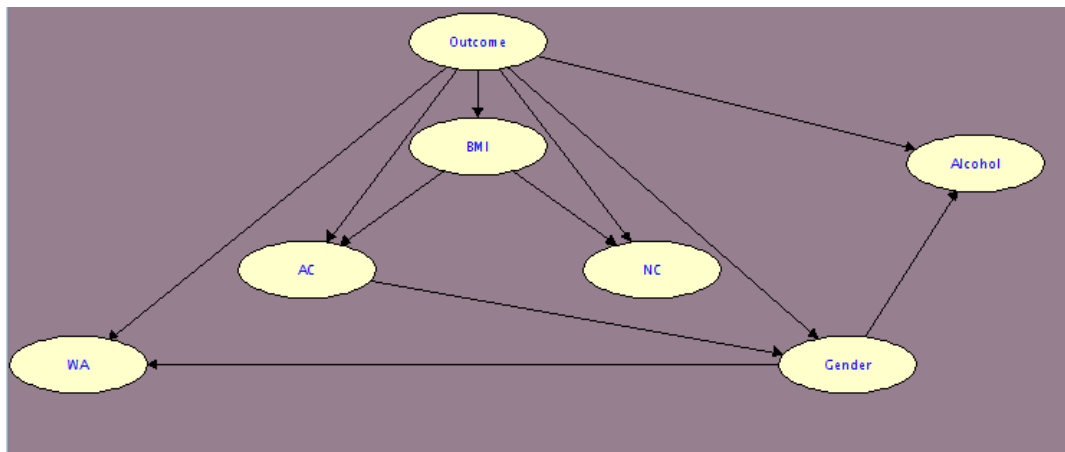


Figure 8: Tree augmented Bayesian network

(WA: Witnessed Apneas, BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference)

One more time, as we saw on NB and LR, we verify high association between AC, NC and BMI. Only three variables have one more dependence beyond the class variable outcome: alcohol and WA are influenced also by gender, as we saw in the network without class information, and gender is influenced by AC.

As we did to NB, we tested different cutoffs to achieve different levels of sensitivity and specificity. For a 100% of sensitivity we choose 2% as cutoff and for 95% of sensitivity we achieve a cutoff of 22%. As internal validation results, the model with 2% cutoff had a sensitivity of 100% and 28% to specificity and using a higher cutoff of 22%, the results were 95% and 38% for sensitivity and specificity respectively (table 8).

On leave-one-out CV, to check the external validation, using the same cutoffs described above, the 2% cutoff has 88% for sensitivity and 23% for specificity while the higher cutoff had 84% of sensitivity and 25% for specificity (table 8).

The marginal probabilities of TAN are presented on figure 9.

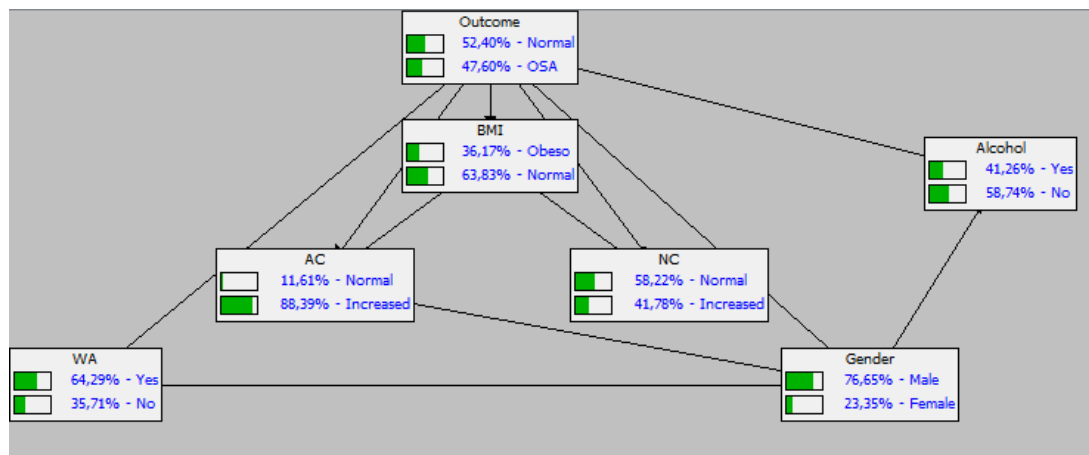


Figure 9: Marginal probabilities of TAN

(WA: Witnessed Apneas, BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference)

5.4. Validation on a second comparable cohort

To test the performance of the model in clinical practice, we collected a second cohort with 33 patients (fig. 10).



Figure 10: Flow diagram for inclusion of patients in the second comparable study

As we expected, there was a higher value for normal results (45%), mainly male gender (76%), with mean age of 53 years (table 7). Of the 18 patients with OSA (54%), 6 (33%) were categorized into mild, 7 (39%) were moderate and 5 (28%) were severe. Statistical testing (table 7) showed that the two cohorts are comparable with respect to the main studied variables (OSA, gender, witnessed apneas, alcohol before sleep, BMI, NC and AC).

Table 7: Characteristics of two samples

	Test N=33	Train N=86	p
OSA, <i>n (%)</i>	18(55)	45(52)	0.833*
Gender, <i>n (%)</i>			0.634*
Male	25(76)	69(80)	
Female	8(24)	17(20)	
Age, <i>mean (sd)</i>	53(14)	56(13)	0.456***
Witnessed apneas, <i>n (%)</i>	21(64)	56(65)	0.469*
Alcohol before sleep, <i>n (%)</i>	17(52)	38(44)	0.466*
BMI, <i>mean (sd)</i>	29(5)	29(4)	0.839***
NC, <i>mean (sd)</i>	41(3)	41(4)	0.879***
AC, <i>mean (sd)</i>	105(12)	105(12)	0.744***
ACr increased, <i>n (%)</i>	31(94)	75(90)	0.723**
NCr increased, <i>n (%)</i>	16(49)	36(43)	0.664*
BMIr Obese, <i>n (%)</i>	12(37)	31(36)	0.956*

*-Chi-square test **-Fischer test ***-t test

OSA: Obstructive Sleep Apnea; BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference; BMIr: Body Mass Index recoded; NCr: Neck circumference recoded; ACr: Abdominal circumference recoded.

5.4.1. Logistic regression model

We applied, using the same cutoffs of the LR model, the same regression equation to the data collected on the second cohort. Using the 10% cutoff, defined for a 100% of sensitivity, the model obtained 100% for sensitivity and 0% for specificity, while using the 25% cutoff, to a 95% of sensitivity, the sensitivity value decreased to 88% but we obtained better specificity (15%). In this last scenario, the model missed two OSA cases: one mild (female, not obese, with PC and NC increased but with no witnessed apneas, with border line result of 5.8 apneas/hour) and one severe (male, not obese, with NC normal, NA increased and a border line result of 33 apneas/hour).

5.4.2. Tree Augmented Bayesian Network and Naïve Bayes

Comparing the results of the two BN models (table 8) aiming 100% for sensitivity, TAN and NB had the same result for sensitivity (94%) on the test sample, but TAN achieved better results on specificity (7%) than NB (0%). The same sensitivity (89%) occurs aiming 95% but TAN (13%) achieved better results for specificity than NB (7%).

Table 8 summarizes the results of the three models results on train and test sample.

Table 8: Results of LR and BN on the train and test sample

	LR			TAN			NB		
	Cutoff	Sens	Spec	Cutoff	Sens	Spec	Cutoff	Sens	Spec
Train									
Baseline, (%)	50	71	73	50	70	73	50	70	73
Aiming 95%, (%)	25	95	35	22	95	38	10	98	33
Aiming 100%, (%)	10	100	5	2	100	28	7	100	25
CV									
Baseline, (%)	50	73	62	50	65	73	50	67	73
Aiming 95%, (%)	25	89	34	22	84	25	10	93	30
Aiming 100%, (%)	10	98	11	2	88	23	7	98	18
Test									
Baseline, (%)	50	56	39	50	72	53	50	56	47
Aiming 95%, (%)	25	88	15	22	89	13	10	89	7
Aiming 100%, (%)	10	100	0	2	94	7	7	94	0

LR: logistic regression; NB: Naïve Bayes; BN: Bayesian networks; TAN: tree augmented Bayesian network, Sens: sensitivity, Spec: specificity; CV: cross-validation.

6. Discussion

The initial step towards defining a decision support system is to assess the need for it. In the two samples, the proportion of normal results was 48% and 45% on train and test sample, respectively, which reveals a large number of normal exams, as we expected, hence supporting the need for a good clinical decision support system to prioritize the waiting list.

Although the large number of clinical variables used (33), only six had significant results in the univariate analysis: body mass index, neck circumference (NC), abdominal circumference, gender, witnessed apneas (WA) and consume of alcohol before sleep. In our study, WA was more frequent in male gender, what is according with the literature. One possible explanation of some authors, referred on background, for the detection of WA, is that it appears more frequent in males, because female bed partners of male patients appear to have a lower threshold for symptom perception and reporting than male bed partners of female patients. Even though in some studies, age appears as a risk factor, in our study it did not end up having significant odd ration (OR) on the univariate analysis. In other hand, these results confirm the inability of some models founded in literature that not considered NC as a risk factor. Because this factor is not described in some articles as important to the development of obstructive sleep apnea (OSA), some physician's don't collect this information on the consultation and our study, one more time, supports the importance of this measure. Even the multivariate logistic regression (LR) model showed that the most discriminative predictive factors for the presence of OSA were the neck circumference and witnessed apneas, which is in accordance with other studies, and supports the inability of some other models that were not build using these factors, in proper prediction of OSA .

The developed models, logistic regression (LR) and Bayesian networks (BN) based, were analyzed considering two different levels for sensitivity. First, models were fitted for 100% sensitivity to avoid false negatives. It was our goal prioritize patients on the waiting list, eventually reduce the number of normal exams performed, but we can't leave patients with moderate or severe OSA without performing polysomnography (PSG). To allow a not so strict level of sensitivity, we experimented also 95%, useful for hospitals with sleep consultation but without sleep laboratory (which have to choose very well patients to refer to other laboratories), which achieved much better results for specificity.

6.1. Overall evaluation

On LR, for 100% sensitivity, we used a cutoff of 10% over the multivariate model's output probability, achieving a sensitivity of 100% and a specificity of 5% in our sample, and 98% and 11% using cross-validation. The values obtained for sensitivity suggests that this approach should evolve towards a clinical decision support system: 100% sensitivity, with high negative predictive value and enough specificity. We achieved a specificity of 35% for a sensitivity of 95% (34% and 89%, respectively, in cross-validation), by choosing a cutoff value of 25%.

When we tested our model in a second comparable cohort, the LR model using the 10% cutoff avoids false negatives but does not reduce the number of normal patients requested to perform PSG. However, this was an expected result given the reduced size of the sample (5% specificity over 15 normal patients represents less than one patient). Using a higher cutoff (25%), the result was good for specificity (15%) but with loss of sensitivity, 88%. Nevertheless, as further explained, we believe that the two OSA patients missed by the model do not hinder the validity of the proposal. First, since the cutoff was optimized for 95% sensitivity, we would expect one OSA patient to be missed (in this case, the mild OSA patient). Then, the severe OSA patient missed by the model actually presented negative values for the two included variables (WA and NC), and could thus be considered an outlier for our modeling.

Considering these results, if we based our clinical decision support system on LR model results, we could foreseen a clinical practice decision support tool with three levels:

1. classify in non-priority group, or avoid requesting PSG, the patients with a probability under 10%;
2. classify in a intermediate level of priority patients in the waiting list for PSG with a probability between 10 and 25%;
3. classify in a priority group patients in the waiting list for PSG with a probability above 25%.

Although a model with only two parameters can be simple to apply in clinical practice, the fact that one of them (witnessed apneas), is not possible to be measured directly (is in fact dependent of partner recognition) or, in some cases, even impossible to determine (if the patient lived alone), might prevent better results of this model's application.

Overall, neck circumference and witnessed apneas information suffices to a logistic regression-based clinical model that can be used to attribute priorities to patients in the waiting lists for PSG, classifying in non-priority group patients in the waiting list with a probability under 25%. At least this model can decrease the number PSG requests, avoiding this exam in the patients with a lower probability.

The significant variables achieved in univariate analysis were used to create the BN using naïve Bayes (NB) and tree augmented Bayesian network (TAN) classifiers. Aiming 100% of sensitivity, we used a cutoff of 7% and 2% for NB and TAN respectively. On internal validation either reveals the same sensitivity, 100%, but TAN had better sensitivity (28%) than NB (25%). The results of sensitivity on leave-one-out CV were better in NB (98%) than in TAN (88%), but TAN had superior specificity results (23%) than NB (18%). If we aim at 95% of sensitivity the cutoff used on NB was 10% and for TAN was 22%. On internal validation, NB had better sensitivity (98%) and worse specificity than TAN (95% and 38%). On CV, NB had better results (93% and 30%) than TAN (84% and 25%).

Based on the results of CV, we expected better results of NB for sensitivity, using both cutoffs and also better specificity aiming 95%, but worse results of specificity than TAN aiming 100%. In fact, when we applied the test sample to our BN models, the results for sensitivity were the same on NB and TAN (94% and 89%) using the two types of cutoff, but the results of specificity were better on TAN (7% and 13%) in the two alternatives when compared with NB (0% and 7%).

As we suggested to the LR model based, we can use these cutoffs results to check priorities to patients recommended to perform PSG using the NB and TAN classifiers:

1. classify in non-priority group, or avoid requesting PSG, patients with a probability under 2% using TAN and 7% using NB;
2. classify in an intermediate level of priority patients in the waiting list for PSG with a probability between 2% and 22% using TAN and between 7% and 10% using NB;
3. classify in a priority group patients in the waiting list for PSG with a probability above 22% using TAN and 10% using NB.

To apply these models on clinical practice we can create a model that uses the two classifiers, NB and TAN. This way, we can create multiple models, based on consensus or not, that compensate the TAN variability with the bias of NB.

6.2. Inference

To explain how the BN can be implemented in a clinical decision support system, we apply the information of one patient tested in various hypothesis of inference. Observed variables are marked red (100%).

Entered a new patient, female, without witnessed apneas nor alcohol consumption before sleep with AC, NC and normal BMI, the probability of OSA result is 1.27% using NB (fig. 11) and higher (21.51%) using TAN (fig.12):

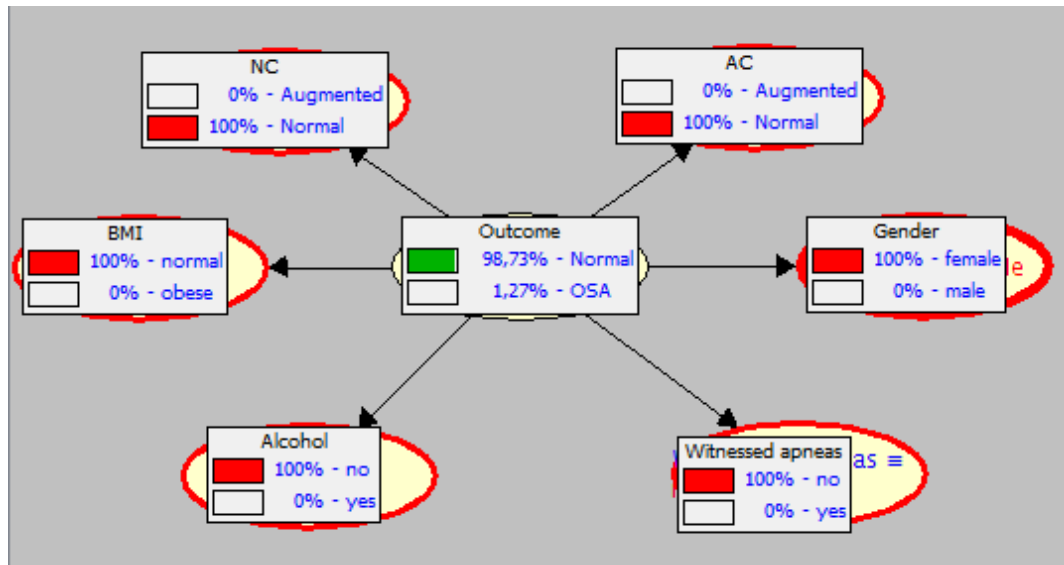


Figure 11: Inference using NB with new evidence: female, without WA nor alcohol consumption before sleep with AC, NC and BMI normal

(WA: Witnessed Apneas, BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference).

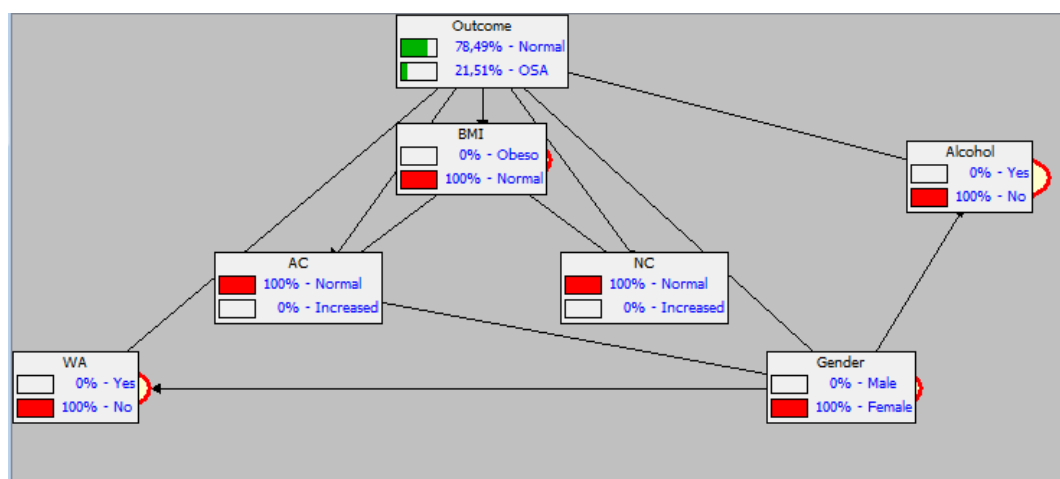


Figure 12: Inference using TAN with new evidence: with new evidence: female, without WA nor alcohol consumption before sleep with AC, NC and BMI normal

(WA: Witnessed Apneas, BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference).

This patient, according to the NB results, had a lower probability of OSA and so, we could consider a non-priority PSG request. The results of PSG confirm this approach, as the result was normal (AHI was 0/h). However, TAN would recommend intermediate level of priority.

Another advantage of BN is the fact that they can deal with missing information. To use the LR model, we need to have access to the two variables present on the regression equation: NC, easy to determine and WA, not so easy, subjective and, in some cases, impossible to achieve.

Imagine that, in the case that we presented above, we don't have access to WA. We can use BN without this information as the probabilities values obtained are based on the learning process that the network made with the information of other patients. The results of NB (fig. 13) showed a probability of 3.26% and 35.92% on TAN.

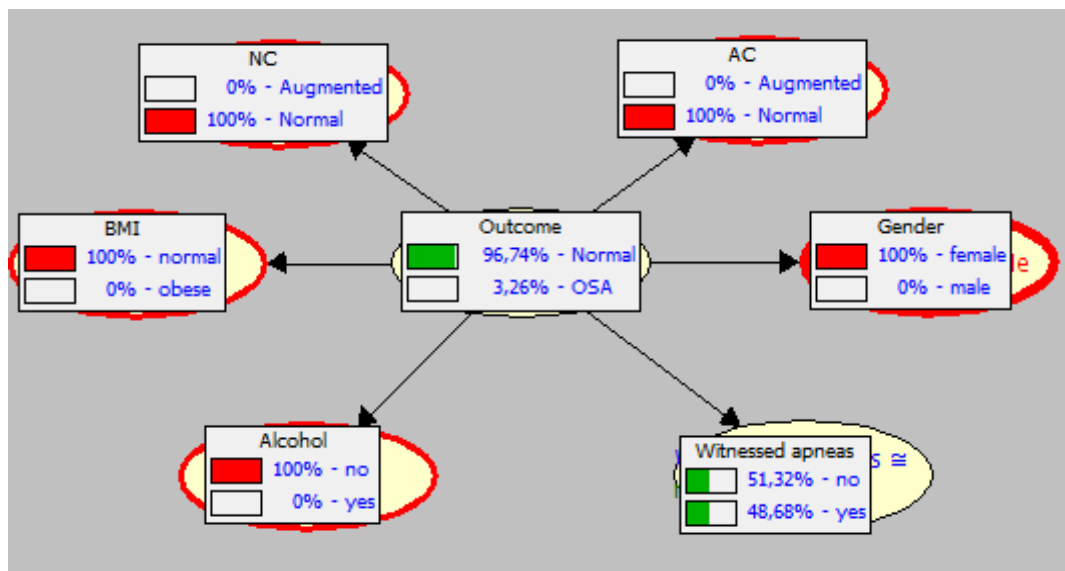


Figure 13: Inference using NB with missing information of WA

(WA: Witnessed Apneas, BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference).

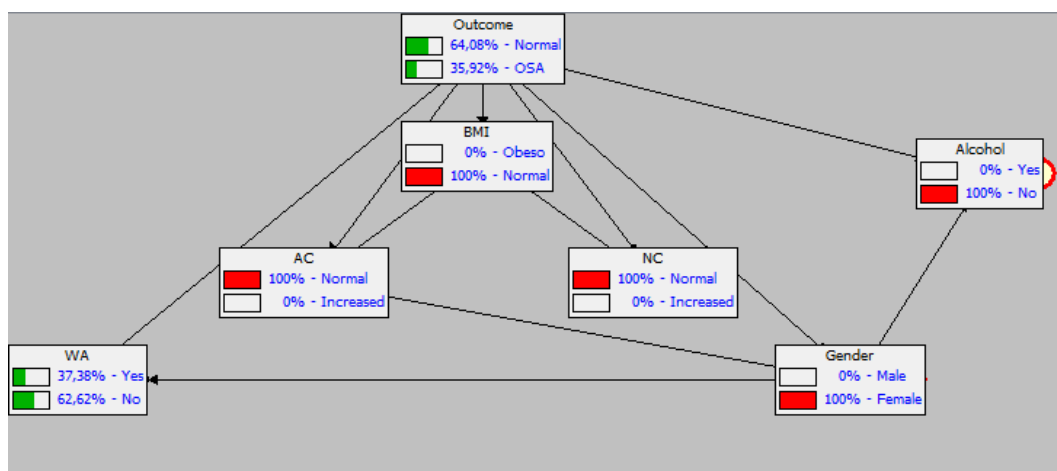


Figure 14: Inference using TAN with missing information of WA

(WA: Witnessed Apneas, BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference).

One more time, NB results were closer to the real result and even without information of WA, the probability of OSA was low.

Suppose that the same patient is recommended to the sleep laboratory to perform PSG by telemedicine or we have access to her electronic registers, but not to the patient, and we want to prioritize based on the information given: only that is a female, with normal BMI and alcohol consumption before sleep. This situation shows the importance of these models, dealing without information of more than one variable, when we have the impossibility of measuring some missing variables (fig. 15 and 16).

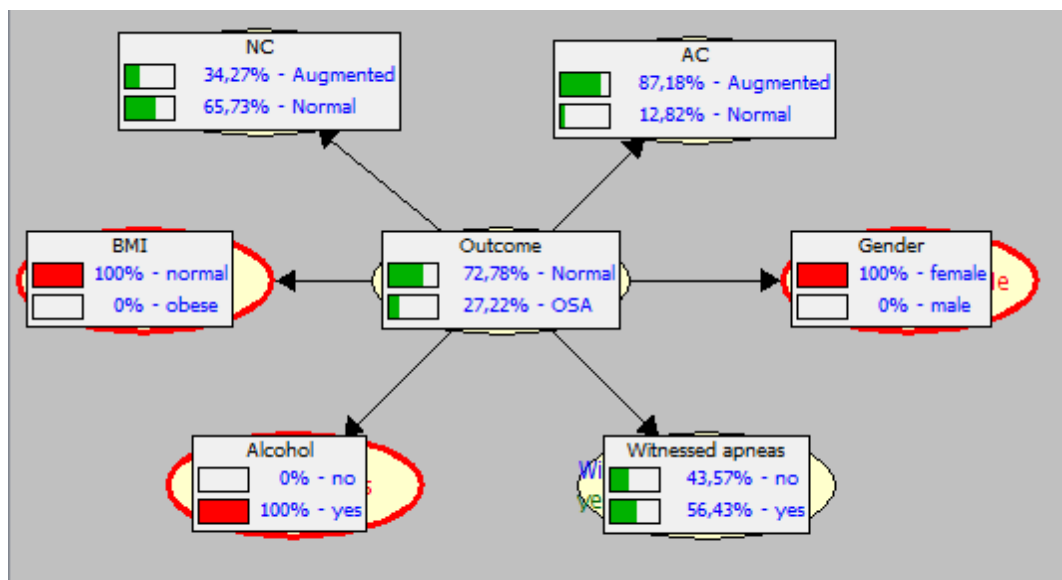


Figure 15: Inference using NB with missing information of WA, AC and NC

(WA: Witnessed Apneas, BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference).

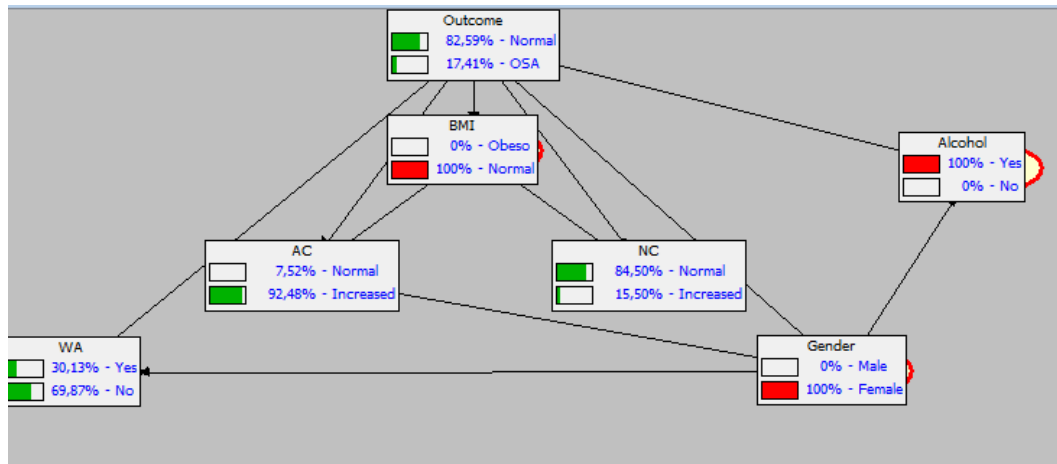


Figure 16: Inference using TAN with missing information of WA, AC and NC

(WA: Witnessed Apneas, BMI: Body Mass Index; NC: Neck circumference; AC: Abdominal circumference).

Comparing these probabilities to the results of figure 13 and 14, we verify that probability for OSA diagnosis decrease using TAN (17.41%) and increase using NB (27.22%) with the lack of information of three variables. If we used the prioritization suggested above using TAN, this patient would be classified as a non-priority group even with only three parameters (gender, alcohol and BMI). These examples shows the bias of NB to classify new cases with missing information and the capability of TAN to deal with these situations as it uses one more variable dependence to classify.

Besides the advantages described above on dealing with missing information, the graphical representation must be seen as another capital gain, mainly the TAN model, that shows more than one dependence between the variables. This is an advantage comparatively to LR based models that don't have this capability.

6.3. Limitations

Some factors were not possible to assess due to the lack of representativeness in the sample (ethnicity, snore, decreased libido, pulmonary hypertension, congestive heart failure, metabolic syndrome, renal failure and hypothyroidism) which may have led to a somewhat biased model. Also, because our study was conducted on patients referred by primary care physicians to the sleep consult, the prevalence of OSA in our sample (52%) was higher than of general populations. Hence, no OSA prevalence estimate can be inferred.

To recode the continuous variables, NC, AC and BMI, we used measures that we found on literature, but there are no standard values to categorized values into normal or altered, so these may lead to some errors in border line characteristics. Here we used value higher than 30 to recode into obese, some authors can

considered 25 (pre-obese). To recode NC and AC we chose the most referred and consensual values criteria on literature but other could exist.

Other question is the metric used to classify OSA. As we explained on background some authors questioned if AHI is the more accurate measure to classify OSA severity. Some suggest the use of RDI in alternative to AHI.

7. Conclusions and recommendations

In this study the main characteristics for obstructive sleep apnea (OSA) were body mass index, neck circumference, abdominal circumference, gender, witnessed apneas and consume of alcohol before sleep.

We used two different techniques to construct the models, one based on logistic regression (LR) and other on Bayesian networks (BN). Using these two techniques, LR and BN, we did not aim to compare the two models directly, but rather show their results to facilitate choices and possibly, complement the two methods. They must be seen as methods that support decisions but do not substitute the physician that has always, according to the clinical history of the patient, the last decision.

With LR, the final model used only two variables on the regression equation: neck circumference and witnessed apneas. The great limitation on the application of this approach is the use of WA, since it is subjective and in some cases impossible to measure.

The great advantages of BN are the fact that they can deal with missing information and the graphical representation that shows not only the values of probabilities given the patient characteristics, but also represents the relationship between variables. This can be an alternative to the traditional statistical measure, odds ratio (OR), that can be interpreted as a relative risk of disease in exposed or not exposed patients.

As we did not find a validated model, tested in Portuguese sleep laboratories, we think that our models consist in a valid method to screen patients with suspicion of OSA, before performing PSG.

The great advantage of our solutions is prioritizing OSA suspicion patients into different levels of priority according to their characteristics, and consequently, their probability of confirming the OSA diagnosis. Other advantage is that the system could manage waiting lists automatically in consultation when the physician inserts patient data.

Eventually, we can reduce the number of normal result exams, optimizing the available resources and making sure that no severe case waits much to time and, consequently, treatment.

8. Future work

Would be interesting to test these models in a multi-center sleep laboratories study to compare our results to others before implementing a decision support system that can be used during pre-polysomnography consultation. This clinical decision support system could be based on multiple models, tested in this work, like logistic regression and Bayesian networks using naïve Bayes and tree augmented Bayesian network.

As the sleep laboratories have also a high number of patients referred by the primary care, we could perform another study to implement a decision support system to help primary care physicians with the decision to send patients to sleep consultation and in this way, prioritize waiting lists for consultation.

9. References

- Ahmadi, N., Chung, S. A., Gibbs, A., & Shapiro, C. M. (2008). The Berlin questionnaire for sleep apnea in a sleep clinic population: relationship to polysomnographic measurement of respiratory disturbance. [Comparative Study]. *Sleep Breath*, 12(1), 39-45.
- Al Lawati, N. M., Patel, S. R., & Ayas, N. T. (2009). Epidemiology, Risk Factors, and Consequences of Obstructive Sleep Apnea and Short Sleep Duration. *Progress in Cardiovascular Diseases*, 51(4), 285-293.
- Aronsky, D., & Haug, P. J. (2000). Automatic identification of patients eligible for a pneumonia guideline. [Evaluation Studies Research Support, Non-U.S. Gov't]. *Proc AMLA Symp*, 12-16.
- Berry, D. (2008). Case study: obstructive sleep apnea. [Case Reports Review]. *Medsurg Nurs*, 17(1), 11-16; quiz 17.
- Bonekat, H. W., & Hardin, K. A. (2003). Severe upper airway obstruction during sleep. [Review]. *Clin Rev Allergy Immunol*, 25(2), 191-210.
- Burnside, E., Rubin, D., & Shachter, R. (2000). A Bayesian network for mammography. [Research Support, U.S. Gov't, P.H.S.]. *Proc AMLA Symp*, 106-110.
- Burnside, E. S. (2005). Bayesian networks: computer-assisted diagnosis support in radiology. [Research Support, Non-U.S. Gov't Review]. *Acad Radiol*, 12(4), 422-430.
- Chung, F., Yegneswaran, B., Liao, P., Chung, S. A., Vairavanathan, S., Islam, S., Shapiro, C. M. (2008). Validation of the Berlin questionnaire and American Society of Anesthesiologists checklist as screening tools for obstructive sleep apnea in surgical patients. [Research Support, Non-U.S. Gov't]. *Anesthesiology*, 108(5), 822-830.
- Collop, N. A., Anderson, W. M., Boehlecke, B., Claman, D., Goldberg, R., Gottlieb, D. J., Schwab, R. (2007). Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med*, 3(7), 737-747.
- Coupé, V. M. H., & van der Gaag, L. C. (2002). Properties of Sensitivity Analysis of Bayesian Belief Networks. *Annals of Mathematics and Artificial Intelligence*, 36(4), 323-356.
- Davies, R. J., Ali, N. J., & Stradling, J. R. (1992). Neck circumference and other clinical features in the diagnosis of the obstructive sleep apnoea syndrome. *Thorax*, 47(2), 101-105.
- Dixon, J. B., Schachter, L. M., & O'Brien, P. E. (2003). Predicting Sleep Apnea and Excessive Day Sleepiness in the Severely Obese*. *Chest*, 123(4), 1134-1141.

- Direcção Geral da Saúde (Plataforma contra a obesidade). Available: <http://www.plataformacontraaobesidade.dgs.pt/PresentationLayer/calculadores.aspx?menuid=161&exmenuid=-1&calc=2.> [Accessed on 27 of May, 2012].
- Doghramji, P. P. (2008). Recognition of obstructive sleep apnea and associated excessive sleepiness in primary care. *J Fam Pract*, 57(8 Suppl), S17-23.
- Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. [Comparative Study Review]. *J Biomed Inform*, 35(5-6), 352-359.
- Epstein, L. J., Kristo, D., Strollo, P. J., Jr., Friedman, N., Malhotra, A., Patil, S. P., Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep, M. (2009). Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. [Practice Guideline]. *J Clin Sleep Med*, 5(3), 263-276.
- Ferber, R., Millman, R., Coppola, M., Fleetham, J., Murray, C. F., Iber, C. (1994). Portable recording in the assessment of obstructive sleep apnea. ASDA standards of practice. *Sleep*, 17(4), 378-392.
- Flemons, W. W., Littner, M. R., Rowley, J. A., Gay, P., Anderson, W. M., Hudgel, D. W., Loube, D. I. (2003). Home diagnosis of sleep apnea: a systematic review of the literature. An evidence review cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society. *Chest*, 124(4), 1543-1579.
- Flemons, W. W., Whitelaw, W. A., Brant, R., & Remmers, J. E. (1994). Likelihood ratios for a sleep apnea clinical prediction rule. *American Journal of Respiratory and Critical Care Medicine*, 150(5 I), 1279-1285.
- Fogel, R. B., & White, D. P. (2000). Obstructive sleep apnea. [Research Support, Non-U.S. Gov't Review]. *Adv Intern Med*, 45, 351-389.
- Gami, A. S., Pressman, G., Caples, S. M., Kanagala, R., Gard, J. J., Davison, D. E., Somers, V. K. (2004). Association of atrial fibrillation and obstructive sleep apnea. [Comparative Study Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. *Circulation*, 110(4), 364-367.
- Gurubhagavatula, I., Maislin, G., & Pack, A. I. (2001). An Algorithm to Stratify Sleep Apnea Risk in a Sleep Disorders Clinic Population. *Am. J. Respir. Crit. Care Med.*, 164(10), 1904-1909.
- Gus, M., Goncalves, S. C., Martinez, D., de Abreu Silva, E. O., Moreira, L. B., Fuchs, S. C., & Fuchs, F. D. (2008). Risk for Obstructive Sleep Apnea by Berlin Questionnaire, but not daytime sleepiness, is associated with resistant hypertension: a case-control study. [Research Support, Non-U.S. Gov't]. *Am J Hypertens*, 21(7), 832-835.
- Hamilton, P. W., Montironi, R., Abmayr, W., Bibbo, M., Anderson, N., Thompson, D., & Bartels, P. H. (1995). Clinical applications of Bayesian belief networks in pathology. [Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S. Review]. *Pathologica*, 87(3), 237-245.
- Hoffstein, V., & Szalai, J. P. (1993). Predictive value of clinical features in diagnosing obstructive sleep apnea. *Sleep*, 16(2), 118-122.
- Iber C, A.-I. S., Chesson AL, Quan SF. . (2007). *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications.* . Westchester: IL: American Academy of Sleep Medicine.

- Jennum, P., & Riha, R. L. (2009). Epidemiology of sleep apnoea/hypopnoea syndrome and sleep-disordered breathing. *Eur Respir J*, 33(4), 907-914.
- Johns, M. W. (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*, 14(6), 540-545.
- Kaimakamis, E., Bratsas, C., Sichletidis, L., Karvounis, C., & Maglaveras, N. (2009, 3-6 Sept. 2009). *Screening of patients with obstructive sleep Apnea syndrome using C4.5 algorithm based on non linear analysis of respiratory signals during sleep*. Paper presented at the Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE.
- Kapur, V. K. (2010). Obstructive sleep apnea: diagnosis, epidemiology, and economics. *Respir Care*, 55(9), 1155-1167.
- Kohler, M. (2009). Risk factors and treatment for obstructive sleep apnea amongst obese children and adults. *Curr Opin Allergy Clin Immunol*, 9(1), 4-9.
- Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in Medicine*, 23(1), 89-109.
- Krieger, A. C., & Redeker, N. S. (2002). Obstructive sleep apnea syndrome: its relationship with hypertension. [Review]. *J Cardiovasc Nurs*, 17(1), 1-11.
- Krug, P. (1999). Snoring and obstructive sleep apnea. [Review]. *AORN J*, 69(4), 792-801.
- Kryger, M. R. T., Dement W. (2005). *Principles and Practices of Sleep Medicine* (Fourth ed.): Elsevier Saunders.
- Kuhlmann, U., Bormann, F. G., & Becker, H. F. (2009). Obstructive sleep apnoea: clinical signs, diagnosis and treatment. [Comment Editorial Review]. *Nephrol Dial Transplant*, 24(1), 8-14.
- Kwiatkowska, M., Atkins, M. S., Ayas, N. T., & Ryan, C. F. (2007). Knowledge-Based Data Analysis: First Step Toward the Creation of Clinical Prediction Rules Using a New Typicality Measure. *Information Technology in Biomedicine, IEEE Transactions on*, 11(6), 651-660.
- Lam, J. C., & Ip, M. S. (2010). Sleep & the metabolic syndrome. [Review]. *Indian J Med Res*, 131, 206-216.
- Lam, J. C., Sharma, S. K., & Lam, B. (2010). Obstructive sleep apnoea: definitions, epidemiology & natural history. *Indian J Med Res*, 131, 165-170.
- Lavrac, N. (2001). Data mining and decision support : a note on the issues of their integration and their relation to expert systems. *Integrating aspects of data mining, decision support and meta-learning (IDDM-2001)*, 1-8.
- Lee, S. M., & Abbott, P. A. (2003). Bayesian networks for knowledge discovery in large datasets: basics for nurse researchers. [Review]. *J Biomed Inform*, 36(4-5), 389-399.
- Lucas, P. (2004). Bayesian analysis, pattern analysis, and data mining in health care. [Review]. *Curr Opin Crit Care*, 10(5), 399-403.
- Lucas, P. J. F., van der Gaag, L. C., & Abu-Hanna, A. (2004). Bayesian networks in biomedicine and health-care. *Artificial Intelligence in Medicine*, 30(3), 201-214.
- Madani, M., & Madani, F. (2009). Epidemiology, pathophysiology, and clinical features of obstructive sleep apnea. *Oral Maxillofac Surg Clin North Am*, 21(4), 369-375.

- Maislin, G., Pack, A. I., Kribbs, N. B., Smith, P. L., Schwartz, A. R., Kline, L. R., Dinges, D. F. (1995). A survey screen for prediction of apnea. *Sleep*, 18(3), 158-166.
- Manber, R., & Armitage, R. (1999). Sex, steroids, and sleep: a review. *Sleep*, 22(5), 540-555.
- Mitchell, T. (1997). *Machine learning*: McGraw Hill.
- Montironi, R., Bartels, P. H., Thompson, D., Scarpelli, M., & Hamilton, P. W. (1995). Prostatic intraepithelial neoplasia (PIN). Performance of Bayesian belief network for diagnosis and grading. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. *J Pathol*, 177(2), 153-162.
- Netzer, N. C., Stoohs, R. A., Netzer, C. M., Clark, K., & Strohl, K. P. (1999). Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S.]. *Ann Intern Med*, 131(7), 485-491.
- Nuckton, T. J., Glidden, D. V., Browner, W. S., & Claman, D. M. (2006). Physical examination: Mallampati score as an independent predictor of obstructive sleep apnea. *Sleep*, 29(7), 903-908.
- Olson, E. J., Park, J. G., & Morgenthaler, T. I. (2005). Obstructive sleep apnea-hypopnea syndrome. [Review]. *Prim Care*, 32(2), 329-359.
- Pagel, J. F. (2008). The burden of obstructive sleep apnea and associated excessive sleepiness. [Review]. *J Fam Pract*, 57(8 Suppl), S3-8.
- Palmer, L. J., & Redline, S. (2003). Genomic approaches to understanding obstructive sleep apnea. [Research Support, U.S. Gov't, P.H.S. Review]. *Respir Physiol Neurobiol*, 135(2-3), 187-205.
- Park, J. G., Ramar, K., & Olson, E. J. (2011). Updates on definition, consequences, and management of obstructive sleep apnea. [Review]. *Mayo Clin Proc*, 86(6), 549-554.
- Petrie, A., & Sabin, C. (2009). *Medical Statistics at a Glance*: Wiley-Blackwell.
- Pillar, G., & Lavie, P. (2011). Obstructive sleep apnea: diagnosis, risk factors, and pathophysiology. [Review]. *Handb Clin Neurol*, 98, 383-399.
- Pouliot, Z., Peters, M., Neufeld, H., & Kryger, M. H. (1997). Using self-reported questionnaire data to prioritize OSA patients for polysomnography. *Sleep*, 20(3), 232-236.
- Practice parameters for the use of portable recording in the assessment of obstructive sleep apnea. Standards of Practice Committee of the American Sleep Disorders Association. (1994). *Sleep*, 17(4), 372-377.
- Punjabi, N. M. (2008). The epidemiology of adult obstructive sleep apnea. *Proc Am Thorac Soc*, 5(2), 136-143.
- Rechtschaffen, A. K. a. A. (1968). A Manual of standardized terminology, techniques and scoring system for sleep stages of human subjects.
- Redline, S., Budhiraja, R., Kapur, V., Marcus, C. L., Mateika, J. H., Mehra, R., Quan, S. F. (2007). The scoring of respiratory events in sleep: reliability and validity. *J Clin Sleep Med*, 3(2), 169-200.

- Rodsutti, J., Hensley, M., Thakkestian, A., D'Este, C., & Attia, J. (2004). A clinical decision rule to prioritize polysomnography in patients with suspected sleep apnea. *Sleep*, 27(4), 694-699.
- Sakellaropoulos, G. C., & Nikiforidis, G. C. (1999). Development of a Bayesian Network for the prognosis of head injuries using graphical model selection techniques. *Methods Inf Med*, 38(1), 37-42.
- Sakellaropoulos, G. C., & Nikiforidis, G. C. (2000). Prognostic performance of two expert systems based on Bayesian belief networks. *Decision Support Systems*, 27(4), 431-442.
- Schurink, C. A., Visscher, S., Lucas, P. J., van Leeuwen, H. J., Buskens, E., Hoff, R. G., Bonten, M. J. (2007). A Bayesian decision-support system for diagnosing ventilator-associated pneumonia. [Research Support, Non-U.S. Gov't]. *Intensive Care Med*, 33(8), 1379-1386.
- Schwab, R. J. (2005). Genetic determinants of upper airway structures that predispose to obstructive sleep apnea. [Research Support, N.I.H., Extramural Research Support, U.S. Gov't, P.H.S. Review]. *Respir Physiol Neurobiol*, 147(2-3), 289-298. doi: 10.1016/j.resp.2005.06.006
- Scutari, M. (2010). Learning Bayesian networks with the bnlearn R package. *Journal of Statistical Software*, 35(3), 1-20.
- Sheldon, A., Belan, I., Neill, J., & Rowland, S. (2009). Nursing assessment of obstructive sleep apnea in hospitalised adults: a review of risk factors and screening tools. [Review]. *Contemp Nurse*, 34(1), 19-33.
- Silber, M. H., Ancoli-Israel, S., Bonnet, M. H., Chokroverty, S., Grigg-Damberger, M. M., Hirshkowitz, M., Iber, C. (2007). The visual scoring of sleep in adults. *J Clin Sleep Med*, 3(2), 121-131.
- Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. (1999). *Sleep*, 22(5), 667-689.
- Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force.* (1999). (1999/08/18 ed. Vol. 22).
- Stierer, T., & Punjabi, N. M. (2005). Demographics and diagnosis of obstructive sleep apnea. *Anesthesiol Clin North America*, 23(3), 405-420.
- Sun, L., Chiu, H.-W., Chuang, C., & Liu, L. (2010). A prediction model based on an artificial intelligence system for moderate to severe obstructive sleep apnea. *Sleep and Breathing*, 1-7.
- Taktak, J., Ajmi Nabli, T., Ben Othmen, H., Mtiraoui, A., & Ben Hadj Hamida, F. (2011). [Therapeutic compliance in primary open angle glaucoma]. *Tunis Med*, 89(2), 142-146.
- Tasali, E., Mokhlesi, B., & Van Cauter, E. (2008). Obstructive sleep apnea and type 2 diabetes: interacting epidemics. [Review]. *Chest*, 133(2), 496-506.
- Tate, J., & Tasota, F. J. (2002). More than a snore: recognizing the danger of sleep apnea. [Case Reports Review]. *Nursing*, 32(8), 46-49.
- Togeiro, S. M., Chaves, C. M., Jr., Palombini, L., Tufik, S., Hora, F., & Nery, L. E. (2010). Evaluation of the upper airway in obstructive sleep apnoea. [Review]. *Indian J Med Res*, 131, 230-235.

- Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. [Comparative Study Research Support, Non-U.S. Gov't Review]. *J Clin Epidemiol*, 49(11), 1225-1231.
- van Gerven, M. A. J., Taal, B. G., & Lucas, P. J. F. (2008). Dynamic Bayesian networks as prognostic models for clinical patient management. *Journal of Biomedical Informatics*, 41(4), 515-529.
- Vaz, A. P., Drummond, M., Mota, P. C., Severo, M., Almeida, J., & Winck, J. C. (2011). Translation of Berlin Questionnaire to Portuguese language and its application in OSA identification in a sleep disordered breathing clinic. *Rev Port Pneumol*, 17(2), 59-65.
- Viner, S., Szalai, J. P., & Hoffstein, V. (1991). Are history and physical examination a good screening test for sleep apnea? *Annals of Internal Medicine*, 115(5), 356-359.
- Weinreich, G., Plein, K., Teschler, T., Resler, J., & Teschler, H. (2006). [Is the Berlin questionnaire an appropriate diagnostic tool for sleep medicine in pneumological rehabilitation?]. *Pneumologie*, 60(12), 737-742.
- Yaggi, H. K., & Strohl, K. P. (2010). Adult obstructive sleep apnea/hypopnea syndrome: definitions, risk factors, and pathogenesis. *Clin Chest Med*, 31(2), 179-186.
- Younes, M. (2003). Contributions of upper airway mechanics and control mechanisms to severity of obstructive apnea. [Research Support, Non-U.S. Gov't Review]. *Am J Respir Crit Care Med*, 168(6), 645-658.
- Young, T., Evans, L., Finn, L., & Palta, M. (1997). Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep*, 20(9), 705-706.
- Young, T., Shahar, E., Nieto, F. J., Redline, S., Newman, A. B., Gottlieb, D. J., for the Sleep Heart Health Study Research Group. (2002). Predictors of Sleep-Disordered Breathing in Community-Dwelling Adults: The Sleep Heart Health Study. *Arch Intern Med*, 162(8), 893-900.
- Young, T., Skatrud, J., & Peppard, P. E. (2004). Risk factors for obstructive sleep apnea in adults. *JAMA*, 291(16), 2013-2016.

Attachments

International Classification of Sleep Disorders

I Dyssomnias

A. Intrinsic Sleep Disorders

1. Psychophysical Insomnia
2. sleep State Misperception
3. Idiopathic Insomnia
4. Narcolepsy
5. Recurrent Hypersomnia
6. Idiopathic Hypersomnia
7. posttraumatic hypersomnia
8. Obstructive Sleep Apnea Syndrome
9. Central Sleep Apnea Syndrome
10. Central Alveolar Hypoventilation Syndrome
11. Periodic Limb Movements Syndrome
12. Restless Legs Syndrome
13. Intrinsic Sleep Disorders NOS

B. Extrinsic sleep disorders

1. Inadequate Sleep Hygiene
2. Environmental Sleep Disorder
3. Altitude Insomnia
4. Adjustment Sleep Disorder
5. Insufficient Sleep Syndrome
6. Limit-Setting Sleep Disorder
7. Sleep-Onset Association Disorder
8. Food Allergy Insomnia
9. Nocturnal Eating (drinking) Syndrome
10. Hypnotic-Dependant Sleep Disorder
11. Stimulant-Dependant Sleep Disorder
12. Alcohol-Dependant Sleep Disorder

13. Toxin-Induced Sleep Disorder
 14. Extrinsic Sleep Disorder NOS
- C. Circadian Rhythm Sleep Disorder
1. Time Zone Change (jet lag) Syndrome
 2. Shift Work Sleep Disorder
 3. Irregular Sleep-Wake Pattern
 4. Delayed Sleep Phase Syndrome
 5. Advanced Sleep Phase Syndrome
 6. Non-24-hour Sleep-Wake Disorder
 7. Circadian Rhythm Sleep Disorder NOS

II Parasomnias

- A. Arousal disorder
1. Confusional Arousal
 2. Sleepwalking
 3. Sleep Terrors
- B. Sleep-Wake Transition Disorders
1. Rhythmic Movement Disorder
 2. Sleep Starts
 3. Sleep Talking
 4. Nocturnal Leg Cramps
- C. Parasomnias Usually Associated With REM Sleep
1. Nightmares
 2. Sleep Paralysis
 3. Impaired Sleep-Related Penile Erections
 4. Sleep-Related Painful Erections
 5. REM Sleep-Related Sinus Arrest

6. REM Sleep-Behaviour Disorder

D. Other Parasomnias

1. Sleep Bruxism
2. Sleep Enuresis
3. Sleep-Related Abnormal Swallowing Syndrome
4. Nocturnal Paroxysmal Dystonia
5. Sudden Unexplained Nocturnal Death Syndrome
6. Primary Snoring
7. Infant Sleep Apnea
8. Congenital Central Hypoventilation Syndrome
9. Sudden Infant Death Syndrome
10. Benign Neonatal Sleep Myoclonus
11. Other Parasomnia NOS

III Sleep Disorders Associated With Medical / Psychiatric Disorders

A. Associated With Mental Disorders

1. Psychoses
2. Mood Disorders
3. Anxiety Disorders
4. Alcoholism

B. Associated With Neurological Disorders

1. Cerebral Degenerative Disorders
2. Dementia
3. Parkinsonism
4. Fatal Familial Insomnia
5. Sleep-Related Epilepsy
6. Electrical Status Epilepticus of Sleep
7. Sleep-Related Headaches

C. Associated With Other Medical Disorders

1. Sleeping Sickness
2. Nocturnal Cardiac Ischemia
3. Chronic Obstructive Pulmonary Disease
4. Sleep-Related Asthma
5. Sleep-Related Gastroesophageal Reflux
6. Peptic Ulcer Disease
7. Fibrositis Syndrome

IV. Proposed Sleep Disorders

1. Short Sleeper
2. Long Sleeper
3. Subwakefulness Syndrome
4. Fragmentary Myoclonus
5. Sleep Hyperhidrosis
6. Menstrual-Associated Sleep Disorder
7. Pregnancy-Associated Sleep Disorder
8. Terrifying Hypnagogic Hallucinations
9. Sleep-Related Neurogenic Tachypnea
10. Sleep-Related Laryngospasm
11. Sleep Choking Syndrome
12. Classification of sleep disorders

SPECIFICATIONS

Digital Specifications for Routine PSG Recordings

Sampling rates	Frequency
EEG	500Hz
EOG	500Hz
EMG	500Hz
ECG	500Hz
Airflow	100Hz
Oximetry	25Hz
Nasal pressure	100Hz
Body position	1Hz
Snoring	500Hz
Rib Cage and Abdominal Movements	100Hz

Filter settings	Low frequency filter	High frequency filter
EEG	0.3Hz	35Hz
EOG	0.3Hz	35Hz
EMG	10Hz	100Hz
ECG	0.3Hz	70Hz
Respiration	0.1Hz	15Hz
Snoring	10Hz	100Hz

Epworth Somnolence Scale (ESS)

How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired?

This refers to your usual way of life in recent times.

Even if you have not done some of these things recently, try to work out how they would have affected you. Use the following scale to choose the most appropriate number for each situation:

0 = Would **NEVER** doze

1 = **SLIGHT** chance of dozing

2 = **MODERATE** chance of dozing

3 = **HIGH** chance of dozing

Situation	Change of dozing
Sitting and reading	
Watching television	
Sitting, inactive in a public place (for example, a theater or a meeting)	
As a passenger in a car for an hour without a break	
Lying down to rest in the afternoon when circumstances permit	
Sitting and talking to someone	
Sitting quietly after a lunch without alcohol	
In a car, while stopped for a few minutes in traffic	

Each question is scored from 0 to 3, giving a maximum score of 24.

CONSENTIMENTO INFORMADO

Liliana Patrícia Pinto Leite, aluna de mestrado em Informática Médica da Faculdade de Medicina da Universidade do Porto pretende realizar investigação, no âmbito da sua dissertação, recolhendo dados de indivíduos sugeridos para realizarem polissonografia no laboratório de estudos do sono do Centro Hospitalar de Vila Nova de Gaia/Espinho, EPE.

- O presente estudo tem como objectivo a construção de vários modelos, construídos quer com base em características dos doentes quer através de ferramentas de data mining e a comparação dos seus resultados de forma a avaliar qual o modelo que obtém maior validade, optimizando a sensibilidade para a predição dos casos mais prováveis de doença.
- A recolha de dados é efectuada uma vez aquando a realização da polissonografia;
- Não estão presentes benefícios ou riscos para o sujeito;
- Será mantida a confidencialidade de todos os dados relativos ao sujeito;
- O sujeito poderá desistir da participação na investigação em qualquer altura, sem ter de dar explicações, apresentar desculpas ou reembolsar despesas;
- O investigador usará de franqueza durante todo o processo, limitando o conhecimento do sujeito aos dados por ele designados como fundamentais face ao objectivo da experiência.

Eu, abaixo assinado _____

declaro que entendo os objectivos, características e duração do estudo e que é de minha livre vontade que participo no mesmo.

____/____/____

Authorization for study realization

Exma. Sr.ª
Liliana Patrícia Pinto Leite

Assunto: *Resposta a pedido de autorização para a realização do estudo "Extração de conhecimento como auxiliar de diagnóstico para o Síndrome de Apneia Obstrutiva do Sono: será possível reduzir o número de polissonografia desnecessárias"*

Informo Vª Ex.ª que o pedido *de autorização para a realização do estudo "Extração de conhecimento como auxiliar de diagnóstico para o Síndrome de Apneia Obstrutiva do Sono: será possível reduzir o número de polissonografia desnecessárias"*, conforme deliberação do Conselho de Administração de 12.Dezembro.2011, está autorizada.

Com os melhores cumprimentos,

Com os melhores cumprimentos,



Júlio Alberto Sampaio
Responsável Serviço
Serviço de Formação Ensino e Investigação

Email jsampaio@chvng.min-saude.pt

Telefone: 227 865 127

Telemóvel: 962053725

www.chvng.min-saude.pt